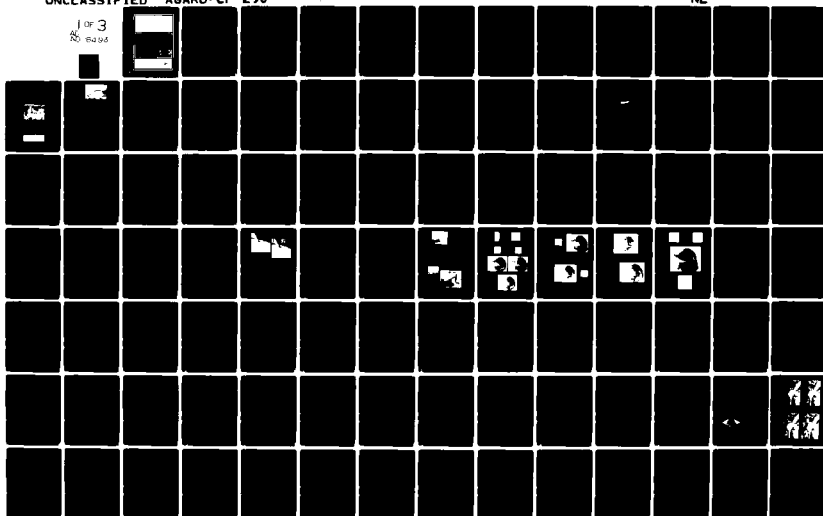AD-A096 493   ADVISORY GROUP FOR AEROSPACE RESEARCH AND DEVELOPMENT--ETC F/G 19/5
              IMAGE AND SENSOR DATA PROCESSING FOR TARGET ACQUISITION AND REC--ETC(U)
              NOV 80
UNCLASSIFIED   AGARD-CP-290   NL

1 OF 3
AD
A096493

LEVEL

AGARD-CP-290

AGARD-CP-290

AGARD

ADVISORY GROUP FOR AEROSPACE RESEARCH & DEVELOPMENT

7 RUE ANCELLE 92200 NEUILLY SUR SEINE FRANCE

AD A096493

AGARD CONFERENCE PROCEEDINGS No. 290

Image and Sensor Data Processing for Target Acquisition and Recognition

DTIC
SELECTED
MAR 1 9 1981
A

NORTH ATLANTIC TREATY ORGANIZATION

DDC FILE COPY

81 3 18 043

AGARD-CP-290

NORTH ATLANTIC TREATY ORGANIZATION

ADVISORY GROUP FOR AEROSPACE RESEARCH AND DEVELOPMENT

(ORGANISATION DU TRAITE DE L'ATLANTIQUE NORD)

AGARD Conference Proceedings No.290

IMAGE AND SENSOR DATA PROCESSING FOR

TARGET ACQUISITION AND RECOGNITION

These Proceedings contain copies of unclassified papers presented at the 40th Technical Meeting,
of the Avionics Panel of AGARD, held in Aalborg, Denmark on 8–12 September, 1980.

THE MISSION OF AGARD

The mission of AGARD is to bring together the leading personalities of the NATO nations in the fields of science and technology relating to aerospace for the following purposes:

– Exchanging of scientific and technical information;

– Continuously stimulating advances in the aerospace sciences relevant to strengthening the common defence posture;

– Improving the co-operation among member nations in aerospace research and development;

– Providing scientific and technical advice and assistance to the North Atlantic Military Committee in the field of aerospace research and development;

– Rendering scientific and technical assistance, as requested, to other NATO bodies and to member nations in connection with research and development problems in the aerospace field;

– Providing assistance to member nations for the purpose of increasing their scientific and technical potential;

– Recommending effective ways for the member nations to use their research and development capabilities for the common benefit of the NATO community.

The highest authority within AGARD is the National Delegates Board consisting of officially appointed senior representatives from each member nation. The mission of AGARD is carried out through the Panels which are composed of experts appointed by the National Delegates, the Consultant and Exchange Programme and the Aerospace Applications Studies Programme. The results of AGARD work are reported to the member nations and the NATO Authorities through the AGARD series of publications of which this is one.

Participation in AGARD activities is by invitation only and is normally limited to citizens of the NATO nations.

THEME

The pace of military engagements is now so rapid that the time taken for target detection and identification is often a serious impediment and limits the overall effectiveness of the response. The problem is becoming relatively more important as the performance of conventional types of sensors are improved, as new sensors are introduced, and as the virtues of using sensors in combination are exploited. It is particularly acute in manned interdictor aircraft and arises in many other avionics applications.

Up to now sensor data has been presented to the operator in essentially "raw" form. Pressures for a more rapid and comprehensive response have resulted in urgent attention to the possibilities of aiding the operator by enhancing visibility of targets to be detected and identified or by accomplishing some or all of the process automatically.

Considerable effort is being directed to accomplish this in many NATO countries. The function of this symposium was to bring together experts in the field to discuss their problems and solutions in a broader context, and to identify: solutions that may have wider applications, theoretical and practical constraints of a general nature, limits of performance that can reasonably be expected in the next generation, and fruitful lines of research and development.

PROGRAM AND MEETING OFFICIALS

CHAIRMAN: Mr R.Voles
Chief Scientist
EMI Electronics Ltd
135 Blyth Road
Hayes, Middx, UB3 1BP
UK

MEMBERS

Mr Y.Brault
Thomson CSF
Division Equipements
Avioniques & Spatiaux
178 Bld Gabriel Péri
92240 Malakoff
France

Dr Ing. M.Vogel
DFVLR e.v.
8031 Oberpfaffenhofen
Post Wessling/obb
Germany

Mr C.W.Cooper
Head Airborne Radar Group
RSRE
St Andrews Road
Malvern, Worcs, WR14 3PS
UK

Dr F.I.Diamond
Technical Director
Communications and Control Division
Rome Air Development Center (AFSC)
Griffiss Air Force Base, N.Y.13340
USA

Dr J.C.Ryles
Chief Scientist
Air Force Avionics Lab/CA
Wright Patterson AFB
Ohio, 45433
USA

AVIONICS PANEL

CHAIRMAN: Dr Ing. M.Vogel
DFVLR e.v.
8031 Oberpfaffenhofen
Post Wessling/obb
Germany

DEPUTY CHAIRMAN: Mr Y.Brault
Thomson CSF
Division Equipements
Avioniques & Spatiaux
178 Bld Gabriel Péri
92240 Malakoff
France

CONTENTS

These Proceedings contain the Unclassified papers presented at the 40th
Technical Meeting of the Avionics Panel of AGARD, which was held in Aalborg,
Denmark, 8–12 September, 1980. It is recommended that the companion
Classified volume be obtained and reviewed at the same time. That volume
contains the discussions.

---

* Not available at time of printing.

† To be found in CP 290 (Supp.) Classified.

---

* Not available at time of printing.
† To be found in CP 290 (Supp.) Classified.

---

* Not available at time of printing.
† To be found in CP 290 (Supp.) Classified.

OVERVIEW

by

R.Voles

As the pace of military operations increases, so the time taken for target detection and identification becomes significant in determining the overall effectiveness of the response. The situation is particularly acute in the manned interdictor aircraft where urgent attention is now being given to enhancing the visibility of targets on the operators display. As the hazards to manned aircraft mount, there is a rapidly growing need for the functions of target acquisition and recognition to be done automatically.

The present Conference has been particularly helpful in giving the participants an overview of the impressive strides that are now being taken in the various areas of this broad technology. Although much remains to be done, the progress to date is very impressive and the user can look forward with confidence to a very effective range of techniques and systems becoming available in the medium term.

As regards the longer term, I shall take the opportunity to indulge in a little speculation into the problems that remain still to be addressed. To do this, I take the human operator as the starting point. The human brain is characterised by 10^{10} neurons each communicating at an effective bandwidth of only 20 Hz but connected to as many as 10^4 neighbours.

Now, in the near future, we shall have VLSI circuits operating at 200 MHz - a factor of 10^7 times faster than the neuron. Although the neurophysiological system contains an impressive 10^{10} neurons, it will not be long before we can match this figure; with densities of 10^6 gates/chip at a cost of £5/chip, the 10^{10} gates would amount to a modest £50k. But despite matching each neuron with a gate operating 10^7 times faster, there is little doubt that an electronic system such as this would be an intellectual moron.

Electronic and neurophysiological systems seem to differ principally in the following respects:

(a) in addition to the electrical process, the neurophysiological system appears to operate a (chemical) memory mechanism;

(b) the neuron has an inherent ability to take an arithmetic sum of up to 10^4 inputs;

(c) the richness of the interconnections between the neurons of the neurophysiological system compared with the extreme simplicity of those in electronic systems.

There is no reason why we should not emulate the neurophysiological processes in (a) and (b) with suitable electronic devices; but the problem of connectivity is another matter. Modern solid-state circuits are essentially two-dimensional structures so even when more connectivity has been introduced (as in the case of the array processor and a custom-designed picture-processing chip) the interconnections are between neighbours in the same plane and usually only amount to about 10 in toto. Thus, the factor of perhaps 10^3 in the richness of the synaptic connectivity of the neuron compared with that of the solid-state circuit seems to endow the neurophysiological system with such a large difference in performance as to put into shadow the adverse speed ratio of 10^7.

These considerations suggest therefore that connectivity is a most important area for research On the theoretical side, it would be of very great value to understand even the rudiments of the principles involved. On the device side, there seems to be a tremendous incentive to invent three-dimensional solid-state structures with a facility for extensive interconnection. The fact that optical devices have an inherent facility for multiple connectivity suggests that they may well have an important part to play in the future.

It is a corollary of these speculations, and an exciting theme with which I can finish, that once we can invent our way into the connectivity problem and have reached the levels achieved by the neurophysiological system, we shall have the capability to make pattern recognition machines with the power of the human brain yet able to operate 10^7 times faster! The future is unimaginably exciting.

THE OPTICAL CONTRAST OF LAND AND SEA TARGETS

by

H.-E. Hoffmann

Deutsche Forschungs- und Versuchsanstalt für Luft- und Raumfahrt e.V. (DFVLR)
Institut für Physik der Atmosphäre

8031 Oberpfaffenhofen
Germany

Summary

In the last years the DFVLR has carried out field experiments for determining the visibility ranges maximum detection range, maximum recognition range or maximum identification range when observing ground to air or air to ground with different observation devices. During visibility tests in Northern Germany in summer and autumn 1977 the inherent contrast of a 1.5 t lorry and a minesweeper-type test boat belonging to the Bundeswehr were also measured. The measurements were taken using a photopic adapted photometer installed in a Bell UH 1D helicopter. Some of these measurement results are presented in diagrams and give information on the influence of the following parameters on the inherent contrast of both the land respectively sea targets: background illumination of target area, angle of elevation, time of day, parts of target, degree of cloud cover, angle of azimuth.

1. Introduction

In the last years field experiments were carried out by DFVLR to determine the ranges in which targets in the air or on the ground can be detected, recognized or identified when using different observation devices. Visibility tests conducted in North Germany during summer and autumn 1977, in the region of Bundeswehr test centre 91 in Meppen and centre 71 in Eckernförde, included measurements of additional optical and atmospheric-optical parameters besides those previously measured. From the ground it had previously been possible to obtain individual test parameters, the horizontal standard visibility and, if required, the adaptation light intensity. In the 1977 tests, using a Bell UH 1D helicopter fitted with a photometer measuring in the visible part of the electromagnetic sprectrum, measurements were also taken of the inherent contrast of objects under observation and of the sky to ground light intensity ratio.

After a short description of the measurements in section 2, section 3 presents results, classified according to particular test parameters for the inherent contrast. These results should provide a basis for some initial general statements, for example concerning the dependence of this parameter on the type of background, type of lighting (shadow, semi-shadow, sunlight), angle of elevation at which the measurements were taken, time of day, angle between direction of measurement and sun's radiation, and degree of cloud cover. It should be noted here that the number of individual values on which these statements are based is in some cases very small, and that it is essential to classify the measured values according to further parameters, e.g. the horizontal standard range of visibility, the sun's angle of elevation.

The values presented here for the inherent contrast should be used to provide a theoretical interpretation of the observation results from [1, 2, 3, 4, 5] by theory of Duntley [6, 7] or may be used as entrance datas for visibility models. During the 1977 field tests initial measurements were taken using the measuring helicopter to determine further optical and atmospheric-optical parameters sky-ground ratio coefficient of reflection, slant standard visibility, atmospheric light, and the contrast between sky and water-horizon. Work was carried out on processing and evaluation of these measurements [8, 9].

2. Measurement of the inherent contrast

The inherent contrast C_0

$$C_0 = \frac{L_0 - L_B}{L_B}$$

(L_0 = light intensity of the object, L_B = light intensity of the background from distances nearly 0)

was calculated using light intensities measured by means of a Pritchard Fotometer installed in a Bell UH 1D helicopter (see Fig. 1). The Fotometer was mounted in the helicopter such that the helicopter vibrations were only transmitted to the Fotometer in a damped form. The Fotometer could be rotated through about 40 degrees in the horizontal plane and about 10 degrees in the vertical plane. The values measured by the Fotometer were registered either on an indicator or by recording apparatus. Values for the inherent contrast involve a measurement error of about

$$\pm 0.02$$

due to the photometer. This error was determined in special measurements.

The fairly wide confidence bands on the mean values (see Chapter 3) were due to the inhomogeneities in light intensities within the photometered objects and their surroundings, and the fact that during measurement the photometer could not be fixed constantly on a particular point of the surface to be photometered.

Measurements were taken with the helicopter between those sections of the tests in which the observation flights were designed to determine the maximum detection range, the maximum recognition range or the maximum identification range [1, 2, 3, 4, 5]. The length of each individual measurement flight for the determination of the inherent contrast and other optical respectively atmospheric-optical parameters was between 20 and 45 minutes.

A minesweeper-type test boat belonging to the Navy was used as the "measurement object" on water; it was also used to obtain the observation values. Fig. 2 shows the test boat and the measurement helicopter during an experiment.

A 1.5 t army lorry was used for the air-ground measurements, and also to obtain the observation values. Fig. 3 shows this lorry in one of its three positions during the Summer 1977 tests.

3. Classification of measurement results according to particular test parameters

The test results are presented in diagrams, classified according to particular test parameters. The diagrams contain the mean values with confidence bands for a statistical probability of 95 %. The number of individual values from which the mean values are taken is shown on the diagrams in brackets after the mean values.

The results for the inherent contrast are sub-divided into results for air-ground and air-water observations. The values for the air-ground inherent contrast are valid for a 1.5 t army lorry (see Fig. 3), and the values for the air-water inherent contrast apply to the Bundeswehr test boat (see Fig. 2).

The diagrams belonging to the following sections 3.1 and 3.2 of this paper are parts of the diagrams of [8]. For more and complete instructions on the results of field experiments 1977 determining values for the optical inherent contrast of a ground and a sea target you must look into the diagrams and tables of this paper.

3.1 Air-ground inherent contrast

The next five figures 4-8 contain values for the inherent contrast of the 1.5 t lorry as target for different measuring conditions: various backgrounds, various lighting conditions, various angles of elevation, and various angles of azimuth between the direction of measurement and the sun's radiation.

Figure 4 gives an impression of the influence of different backgrounds on the inherent contrast. These measurements were taken in front of "green grass and the edge of a wood", "green cornfield", and "yellow sand". The time of measurement was the middle of June. During the tests of figure 4 the target was in shadow. Then the value for the inherent contrast for the "green grass and edge of wood" background was -0.28, for the "green cornfield" -0.46 and for "yellow sand" -0.71.

Also the type of light in the target area altered the inherent contrast of the 1.5 t lorry. If this was in front of the wood (see figure 5), then the inherent contrast was -0.28 for shadow, -0.33 for semi shadow and -0.42 for sunshine.

The next figure 6 shows the influence of different angles of elevation when measuring the inherent contrast of this specific target on the ground. For example with a background of "green grass and edge of wood" in shadow, then the absolute value of the inherent contrast was decreasing with increasing angle of elevation. At angle of 6° it was -0.34, at an angle of 12° it was -0.29 and at an angle of 18° it was -0.23.

Only six individual values for forming mean values were available to give information in which manner the inherent contrast is changing when time of day and with that the angle between the direction of measurement and the sun's radiation are changing. In the specific case of figure 7 - the angle was changing from 165° to 240° - the inherent contrast had changed from -0.34 to -0.18.

When measuring from an elevation angle of 18° the specific ground target used during the measurement presented two surfaced of more or less equal area but very different brightness (see figure 3). The values of the inherent contrast for the darker side panel were -0.29 and for the brighter roof +0.14 (see figure 8).

3.2 Air-sea inherent contrast

When measuring the inherent contrast of a minesweeper-type test boat the influence of the following parameters was tried to determine: various degrees of water brightness, various degress of water and boat brightness and various angles of azimuth between the direction of measurement and the sun's direction.

When observing in approximately the direction of the sun's radiation and the sun is shining then direct sun's radiation reflected on the water or reflected sky light only can be the background. The inherent contrast value was -0.67 when the reflected sky light only was the background and the inherent contrast value was -0.93 when the reflected direct sun radiation was the background (see figure 9). These values are only valid for the test conditions belonging to figure 9.

Because of the dependence of the reflectivity of water the brightness of water background is different in front of and behind a ship. The influence of these various brightnesses on the inherent contrast shows figure 10. The two values of figure 10, -0.54 water background behind the boat and -0.43 water background in front of the boat - were determined when the angle of azimuth between the direction of measurement and the sun's radiation was 1o1° and the degree of cloud cover was 8/8.

The plotted mean values for the inherent contrast of figure 11 were got when the brightnesses of the water and of the boat were changed by variation of the cloud covery. When the cloud cover was 0/8 - 3/8 then the inherent contrast value was -0.72 and at a cloud cover degree 6/8 - 8/8 then the inherent contrast value was -0.59. During these experiments the angle of azimuth between the direction of measurement and the sun's radiation was approximately 180°. For water brightness ever reflected sky light only was measured.

The last two figure 12 and 13 give informations on inherent contrast values when observing a ship from different directions. The first values of these figures were measured when direction of measurement and direction of the sun's radiation were approximately coincident. That means the boat was brighter than its background. The second values represent inherent contrast values when direction of measurement and direction of the sun's radiation were approximately opposite. That means the boat was darker than its background.

With an angle of azimuth of approximately 0° between the direction of measurement and the sun's radiation the inherent contrast showed positive values: +0.83 for degree of cloud cover between 0/8 and 3/8, +0,76 for degree of cloud cover between 6/8 and 8/8.

With an angle of azimuth of approximately 180° between the direction of measurement and the sun's radiation the inherent contrast showed negative values: -0.72 for degree of cloud cover between 0/8 and 3/8, -0.59 for degree of cloud cover between 6/8 and 8/8.

References

[1] Hoffmann, H.-E.
    Buell, R.H.
Die maximale Auffaßreichweite und die maximale Erkennbarkeitsentfernung bei Beobachtung von Bodenzielen aus dem Hubschrauber,
DFVLR IB, 1976.

[2] Hoffmann, H.-E.
    Buell, R.H.
Die maximale Auffaßreichweite bei Beobachtung Luft - Land und Luft - Wasser mit Lichtverstärkergeräten,
DFVLR IB, 1978.

[3] Hoffmann, H.-E.
    Buell, R.H.
    Kuehnemann,W.
Die maximale Auffaßreichweite und die maximale Erkennbarkeitsentfernung bei Beobachtung von Bodenzielen aus dem Hubschrauber
1. Fortsetzung
DFVLR IB, 1978

[4] Hoffmann, H.-E.
    Buell,R.H
    Kuehnemann,W.
Die maximale Auffaßreichweite und die maximale Identifizierungsentfernung Luft - Wasser und die maximale Auffaßreichweite Wasser-Luft
DFVLR IB,1978

[5] Hoffmann, H.-E.
    Buell,R.H.
    Kuehnemann,W.
Summaries of some reports of 1978 concerning results of field experiments for visibility
DFVLR-Mitt. 79-11, 1979

[6] Middleton,W.E.K.
Vision through the atmosphere
Toronto, University of Toronto Press, 1963

[7] Duntley,S.Q.
The reduction of apparent contrast by the atmosphere
Josa 38, 1948

2-4

[8] Hoffmann, H.-E.

Optical and atmospheric optical
parameters during visibility
tests in the north of Germany 1977
European Space Agency
ESA-TT-615,1980

[9] Hoffmann, H.-E.

Some results of preliminary tests
on values of sky-sea contrast (visible
horizon) as a function of environmental
parameters
European Space Agency
ESA-TT-623, 1980
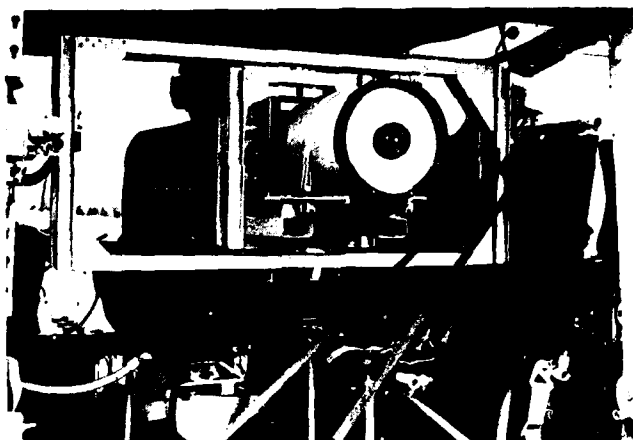
Fig. 1 Photometer in Bell UH helicopter

Fig. 2 Minesweeper-type test boat with
measurement helicopter
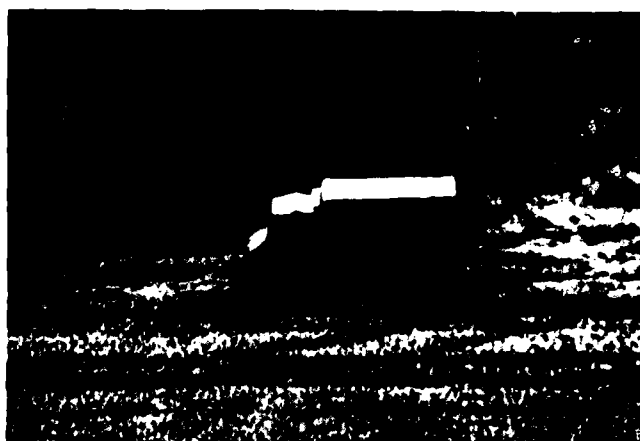
Fig. 3 1.5t army lorry at the edge of a wood
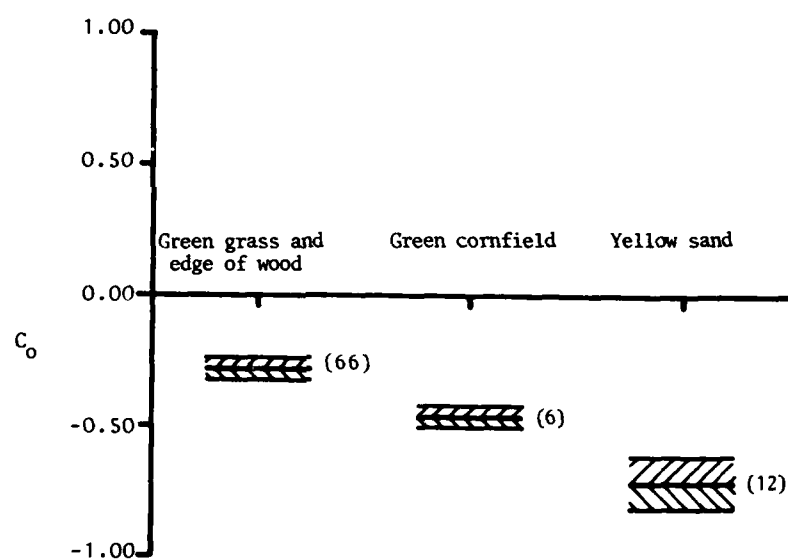
Fig.4 Inherent contrast C_o for various backgrounds.
Target area in shadow
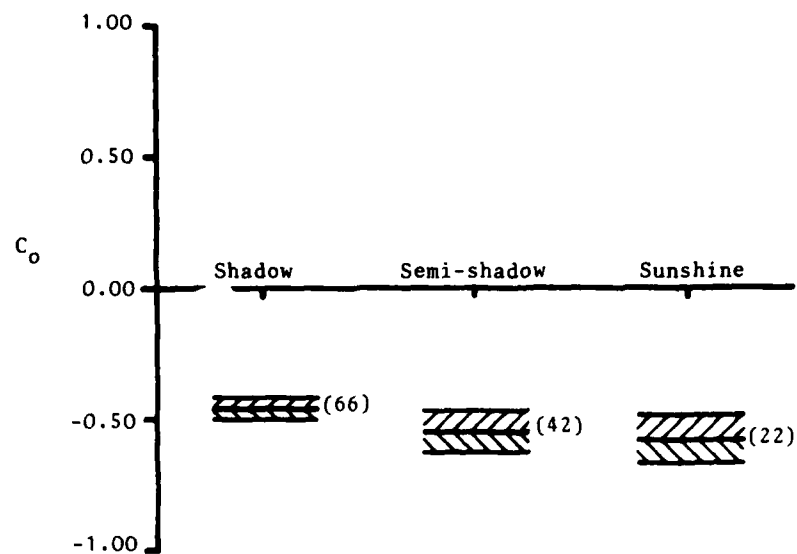
Fig.5 Inherent contrast C_o for various lighting conditions.
Target area: Green grass and edge of wood.
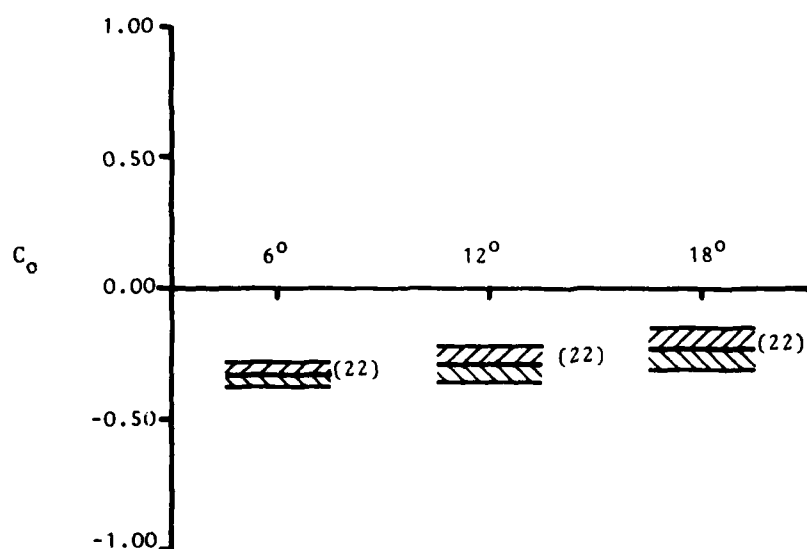
Fig.6 Inherent contrast C_o for various angles of elevation.
Terget area: Green grass and edge of wood in shadow
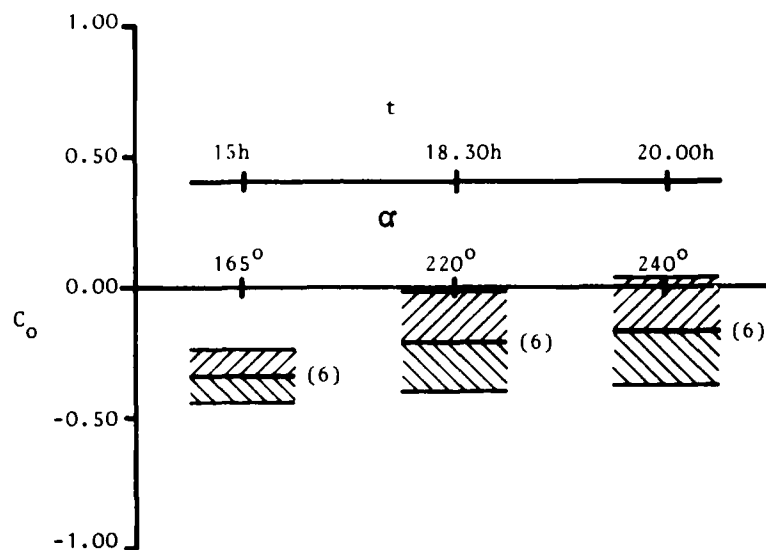
Fig.7 Inherent contrast C_o for various times of day t and angles α between the direction of measurement and the sun's radiation.

Target area: Green grass and edge of wood in shadow
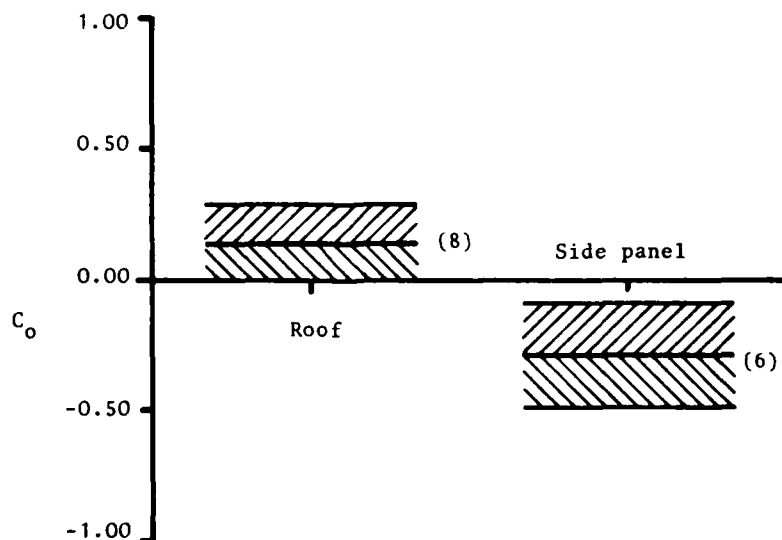
Fig.8 Inherent contrast C_o for varous parts of the 1.5t lorry.

Target area: Green grass and edge of wood in sunshine.
Angle of elevation: 18°
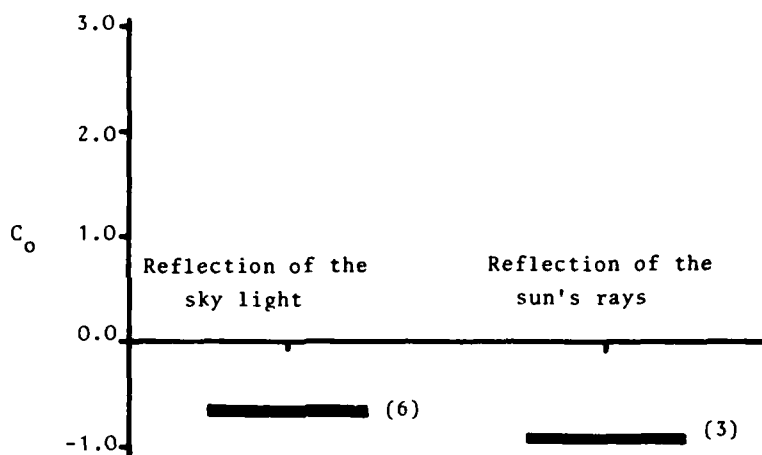
Fig.9   Inherent contrast C_o for reflection of sky light
and reflection of the sun's rays.

Angle of elevation: 5°
Degree of cloud cover:3/8 - 4/8
Angle of azimuth between direction of measurement
and the sun's radiation: 160°
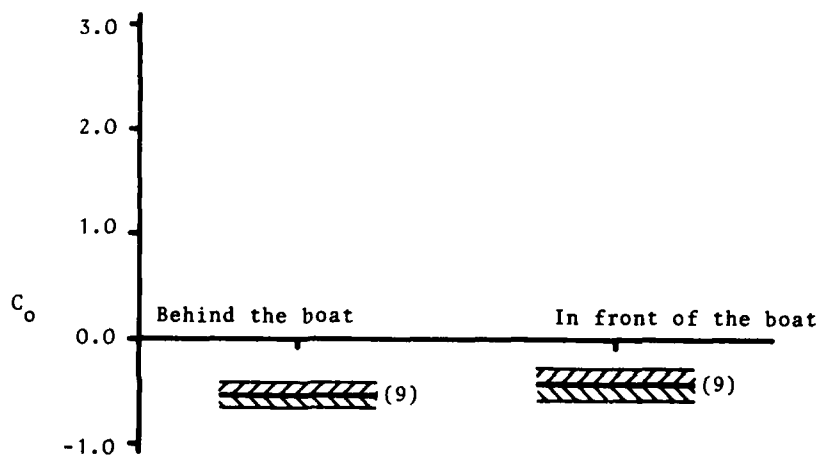
Fig.10   Inherent contrast C_o for reflection of the sky
in front of and behind the boat.

Degree of cloud cover: 8/8
Angle of azimuth between the direction of measure-
ments and the sun's radiation: 101°
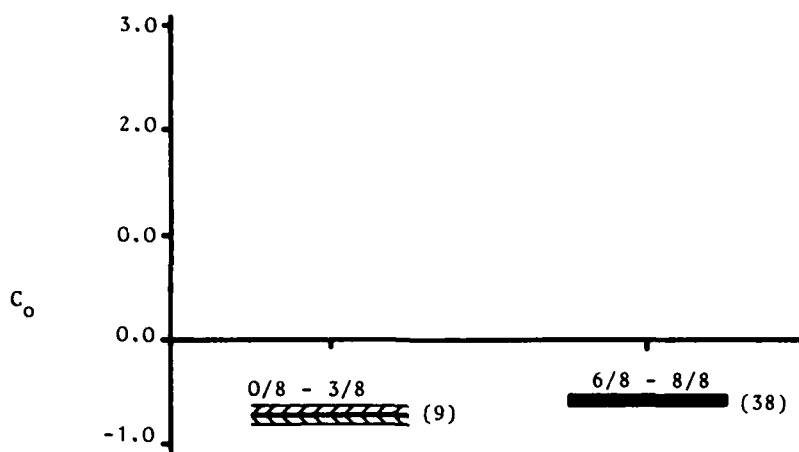
Fig.11 Inherent contrast C_o for various degrees of cloud
        cover.

        Angle of azimuth between the direction of measure-
        ment and the sun's radiation: $180^o \pm 30^o$
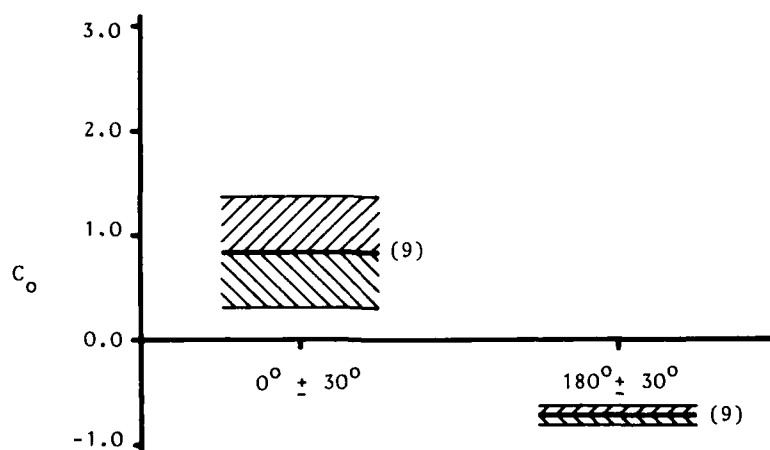
Fig.12 Inherent contrast C_o for various angles of azimuth
        between the direction of measurement and the sun's
        radiation.

        Degree of cloud cover: 0/8 - 3/8

$$C_o$$
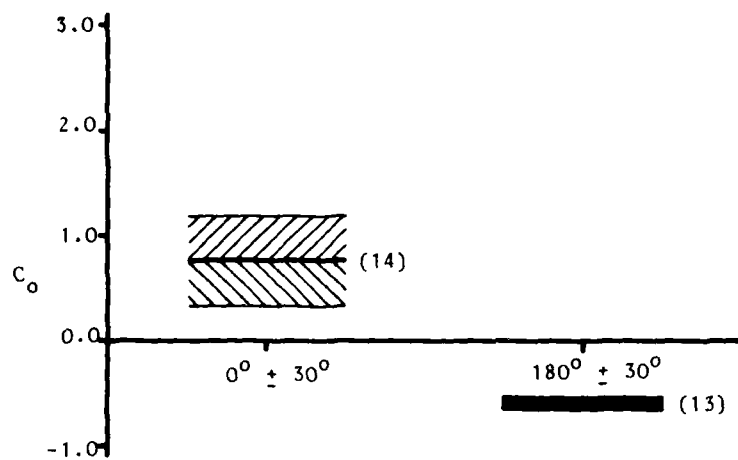
Fig.13   Inherent contrast C_o for various angles of azimuth between the direction of measurement and the sun's radiation.

Degree of cloud cover: 6/8 - 8/8

AN APPROACH TO AIRBORNE INFRARED SEARCH SET PERFORMANCE MODELING

Nancy E. MacMeekin
Naval Air Development Center
Warminster, Pennsylvania 18974 U.S.A.

ABSTRACT

An airborne search set mathematical model is described which predicts sensor performance in terms of probabilities of detection and false alarm in addition to calculating signals available for processing by the sensor. Inputs and submodels which describe and/or calculate target signatures, backgrounds, atmospheric effects and sensor design are discussed. Sensor performance is illustrated for several atmospheric conditions.

INTRODUCTION

Naval aircraft and surface vessels are subject to attack by high-speed enemy aircraft and missiles launched from ranges of the order of 100 kilometers. The task of detecting and tracking such threats at ranges sufficient to permit neutralizing their effects is being addressed in a number of investigations. One approach is that of an infrared search and track set, which could passively acquire and track missiles and aircraft at long ranges.

In this paper, I shall describe this search and track concept, discuss the mathematical modeling approach being pursued at the Naval Air Development Center to evaluate this concept, and present some examples of the results of the modeling.

CONCEPT DEFINITION

An airborne infrared search and track set is defined as an aircraft subsystem which can passively detect, localize, and track a target within a prescribed search volume by sensing the infrared radiation emitted and reflected by the target. Such a system is required to search a wide field of view, often 360° in azimuth, and to detect and track point targets at ranges of tens or hundreds of kilometers; consequently, the target signals may be quite weak and easily lost in background clutter. This kind of set is different from an airborne threat warning receiver in that the warning receiver is designed to warn a pilot of imminent danger, generally from the rear, within 10 kilometers, to which he must respond immediately by evasive maneuvering, decoy deployment, or other countermeasures. For a warning receiver, then, the field of view is relatively small, perhaps 40° x 60°, and the signal levels high.

A concept diagram of a simple airborne infrared search and track set is shown in Figure 1. A sensor enclosure, rotatable through 360°, houses the detectors, optics, cooler, and electronics comprising the sensor. The gimballed sensor is stabilized so that the aircraft's random motion does not affect the search of a given volume of space. As the sensor assembly scans in both azimuth and elevation, infrared radiation passes through the protective window and, by means of the optics, is focused on the cryogenically cooled detectors. The detectors convert changes in incident infrared radiation in selected filtered wavebands to signal voltages. The signal processor in the sensor enclosure amplifies the low-level signals from the detectors and then multiplexes the signals by sampling each detector channel and combining the data into one channel. The multiplexed signal leaves the rotating sensor enclosure through slip rings and enters a signal processor/conditioner where various techniques can be used to separate the target from background clutter caused by things such as cloud edges or sun glints. A computer generates synthetic video for the pilot and provides tracking information to the weapons system.

The application of this concept to naval operations is illustrated in Figure 2. An airborne infrared search and track set could be used for defending the fleet against attacking enemy aircraft and against anti-ship missiles launched from distances beyond the ship's horizon. An equipment of this type could also be used for aircraft self-defense, whether to alert a fighter to take necessary action or to alert a patrol aircraft in time to enable it to get out of enemy interceptor range. Because the infrared search and track set is passive, it could be of particular value in situations in which radiation emission control has been mandated. In addition, it has advantages against threats of small radar cross-section and where radar jamming is in effect.

MODEL DESCRIPTION

The objective of the modeling investigations at the Naval Air Development Center is to explore the utility of an airborne infrared search and track set in meeting operational needs. At issue is whether a system of desired performance characteristics can be built within technological, cost, size and weight constraints. The critical problem seems to be detecting the target with an acceptable false alarm rate, rather than tracking it; consequently, the modeling effort thus far has not included the tracking functions.

The approach taken to modeling performance of airborne infrared search sets has been to acquire a model that includes the necessary attributes, but that is not too expensive and complicated to run. The source of our model, described in reference 1, is B-K Dynamics, Inc. Table I summarizes the kinds of data required to run the model and the calculations produced by the model.

TABLE I

AIRBORNE INFRARED SEARCH SET MODEL: INPUTS AND OUTPUTS	
Inputs	**Outputs**
Target Signature	Contrast Irradiance
Background Characteristics	Probability of Detection
Atmospheric Transmission and Emission	Probability of False Alarm
Operational Geometry	
Sensor Characteristics	
Processing Characteristics	

The infrared signature of an aircraft or missile comprises varying amounts of radiation from the exhaust plume, radiation from the aerodynamically heated fuselage, and reflected radiation. The proportion of each of these is determined by the environmental conditions and by the target's configuration, type of propulsion system and fuel, thrust, airspeed, and aspect angle relative to the sensor. Although the plume radiation might be expected to dominate the signature, the plume can be totally or nearly occluded by the fuselage of the target. The model can calculate a simple target signature, or it can accept externally generated target data.

The backgrounds that are encountered in open-ocean operations are sun, sea, sky, and clouds. The most simple to consider are clear-sky backgrounds, which can be described rather well in terms of the air temperature and the zenith angle. The radiation from the sea is a function of its own thermal emission and the reflected radiation from the sky, both of which are dependent upon the sea state and the angle at which the sea is viewed. Backgrounds are complicated by the presence of sunglints from the sea, the presence of clouds of nonuniform radiance, and perhaps man-made objects which are not threats. In the model, the background is described in terms of an array of 200 small solid angular elements, each of whose emissivity, temperature, and range from the sensor may be specified by the user. The model could be modified to accept a tape of background data.

The atmosphere absorbs and emits radiation in amounts that depend upon the altitude, length, temperature, water vapor and aerosol content of the path, and upon the wavelength. In the model, the LOWTRAN 4 atmospheric model developed by the Air Force Geophysical Laboratory is used to calculate atmospheric transmission and emission, as described in reference 2. We are trying to improve the modeling by including cloud-free line of sight data and by inserting into the atmospheric model meteorological data for specific geographical areas and seasons.

When the user has specified the operational geometry in terms of sensor and target altitudes, range, and target aspect angle, the model has all the information it needs to calculate the irradiances at the sensor aperture produced by the target and the background.

In the model, the sensor is described in terms of the spectral bandpasses of the detectors and the filters, the size of the aperture, the optical efficiency, the size, spacing, and material of the detectors, the scan rate, the sampling time, and the characteristics of the preamplifiers and electronic filtering. The model uses the spectral response characteristics of the optical filter and the detectors to integrate the target and background irradiances and calculate the contrast irradiance, that is, the difference in aperture irradiance between when a target is viewed and when it is not. The contrast irradiance may be understood as the signal available for the sensor to detect. The model proceeds to calculate at each sampling time the power incident on each detector from which signal voltages and voltage variances are derived.

Finally, the user specifies his/her choice of threshold processing, and the model calculates the probability of detection for each sample when the target is in the scene, and the probability of false alarm when the target is not in the scene. Gaussian noise statistics are assumed. The threshold can be specified directly, as a voltage, or it can be calculated by the model relative to some user-preselected multiple of the sensor's noise equivalent irradiance; alternatively, the threshold can be specified in terms of a desired false alarm rate, or, if desired, a specified technique of adaptive thresholding can be employed.

EXAMPLES OF MODELING RESULTS

A sample problem has been chosen to illustrate the use of the model. An aircraft is traveling at high speed, with its afterburner on. It is at an altitude of 11 kilometers and is being viewed against a clear sky background from a horizontal aspect angle of 20° relative to the nose-on direction. The sensor is at long ranges, at an altitude of 5 kilometers. It is assumed that a waveband in the 3.5 to 5.2 micrometer region is of interest.

Figure 3 illustrates the spectral radiant intensity of the aircraft. Figure 4 is a plot of the spectral transmittance through a tropical atmosphere for two different path lengths between the two altitudes. The effect of carbon dioxide absorption in the 4.2 to 4.5 micrometer band is clearly seen. Figure 5 shows, for a fixed path length of 100 kilometers, the dependence of the atmospheric transmittance on the particular climatic model used. The differences would be larger at lower altitudes. Figure 6 illustrates the apparent spectral radiant intensity of the target seen at various ranges. Because most of the plume radiation falls in the region of strong carbon dioxide absorption, the target's signature is severely attenuated.

The contrast irradiance calculated for the sample problem is shown as a function of range in Figure 7 for three different wavebands. The wavebands selected are the broad band, and a band on either side of the carbon dioxide absorption. The noise equivalent irradiance (NEI) of an example 3.5 to 5.2 micrometer sensor is also shown on the figure. By assuming a threshold value of a multiple of NEI, one can find detection range by finding the intersection of the contrast irradiance for the 3.5 to 5.2 micrometer waveband with the threshold value.

In Figure 8, the probability of detection of the aircraft at a range of 100 kilometers against the clear sky background and two probability of false alarm curves have been plotted as functions of the threshold setting. One of the false alarm curves represents the probability of threshold crossing when only clear sky is in the field of view. The other represents the likelihood of threshold crossing when a cloud is in the field of view. The threshold is expressed as a multiple of NEI. The optimum threshold is seen to be dependent on the acceptable probability of false alarm, in addition to the required probability of detection.

CONCLUSION

I have described the model being used at the Naval Air Development Center to calculate the performance of postulated airborne infrared search sets. Results of sample calculations have been presented to illustrate some of the effects of varying target, atmospheric, sensor and processing characteristics. It appears that this model, which is relatively simple and inexpensive to run, will be quite useful in qualitative and quantitative assessments of the feasibility of building an airborne infrared search set to meet US Navy requirements.

REFERENCES

1. N. J. Chesser, B-K Dynamics, Inc, Rockville, MD, "E-O Simulation Computer Code (EOSIM) - Final Report," 1979, Naval Research Laboratory Contract No. N00014-77-C-0668.
2. J. E. A. Selby et al., Air Force Geophysical Laboratory, Hanscom Air Force Base, MA, "Atmospheric Transmittance/Radiance: Computer Code LOWTRAN 4," 1978, Report No. AFGL-TR-78-0053.
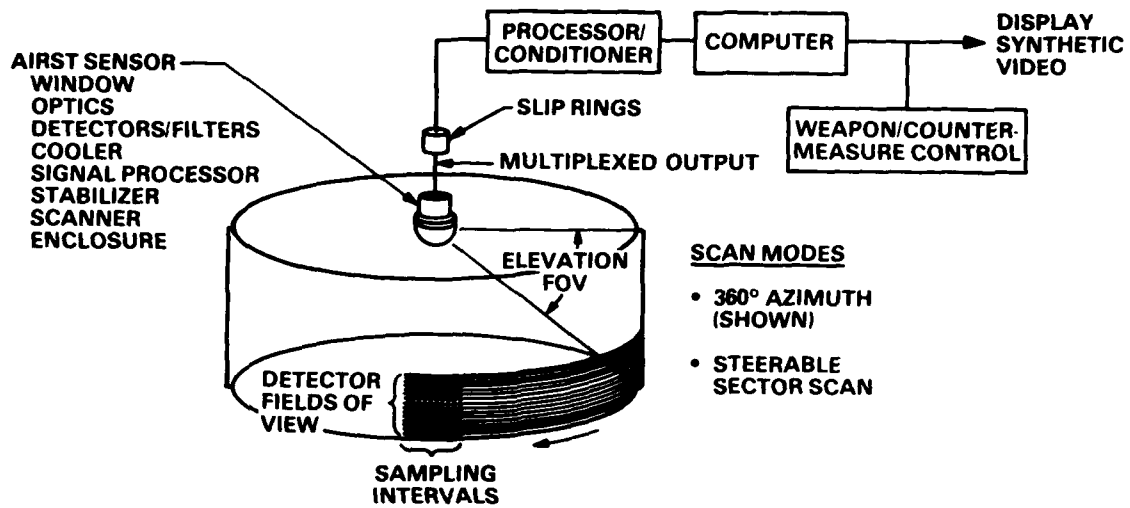
Figure 1. Concept drawing of an infrared search and track set.
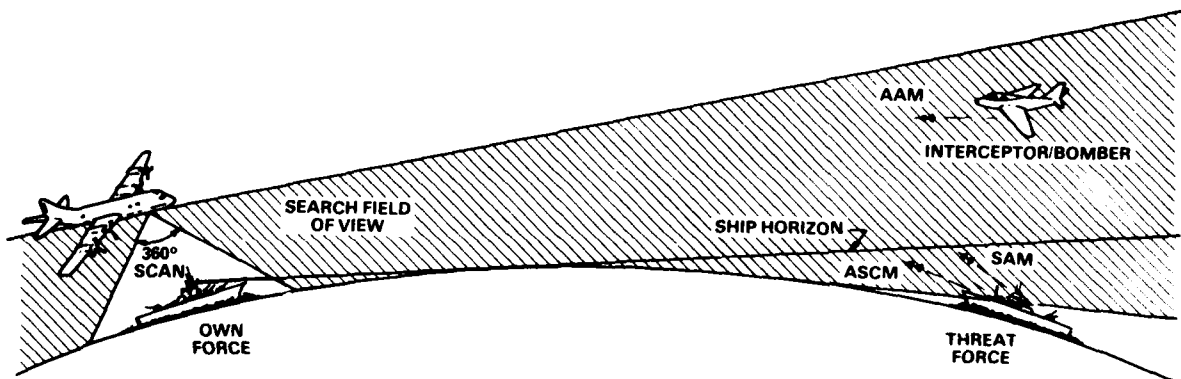
Figure 2. Utility of an airborne infrared search and track set for naval operations.

WAVELENGTH (μm)
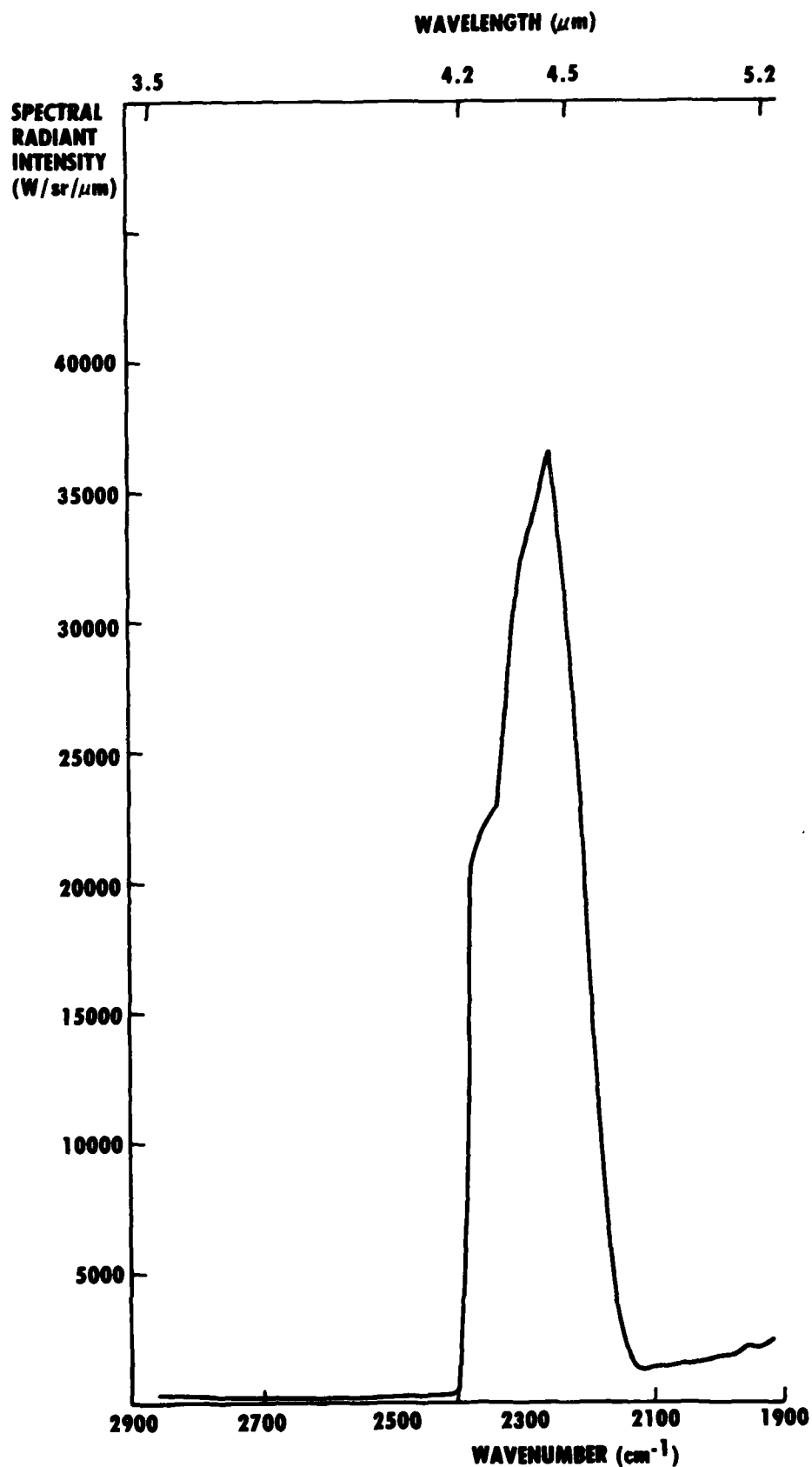
Figure 3. Spectral radiant intensity of an afterburning aircraft.
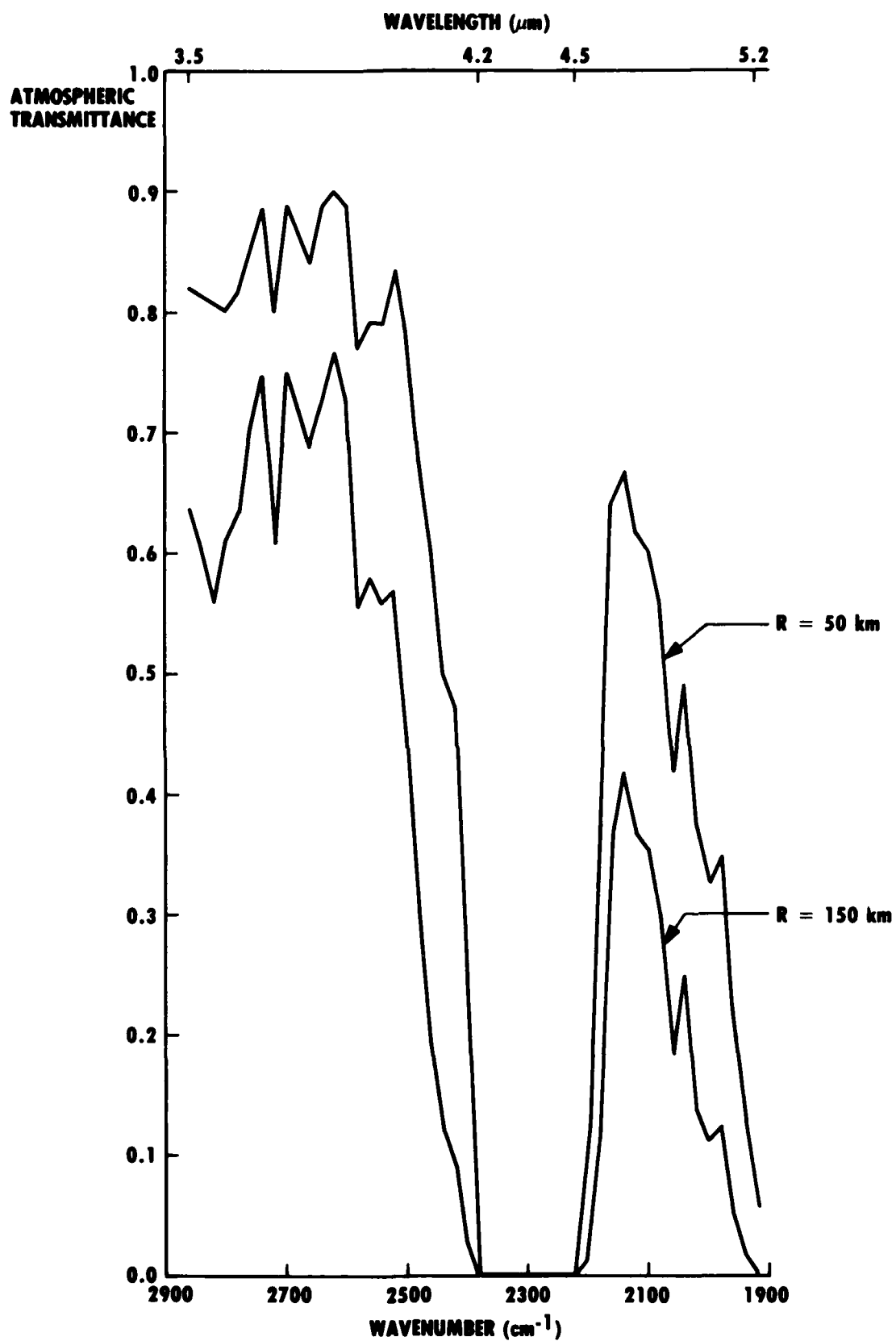
Figure 4. Atmospheric transmittance for two paths in a tropical atmosphere. The observer is at 5 km altitude; the other end of the path is at 11 km.
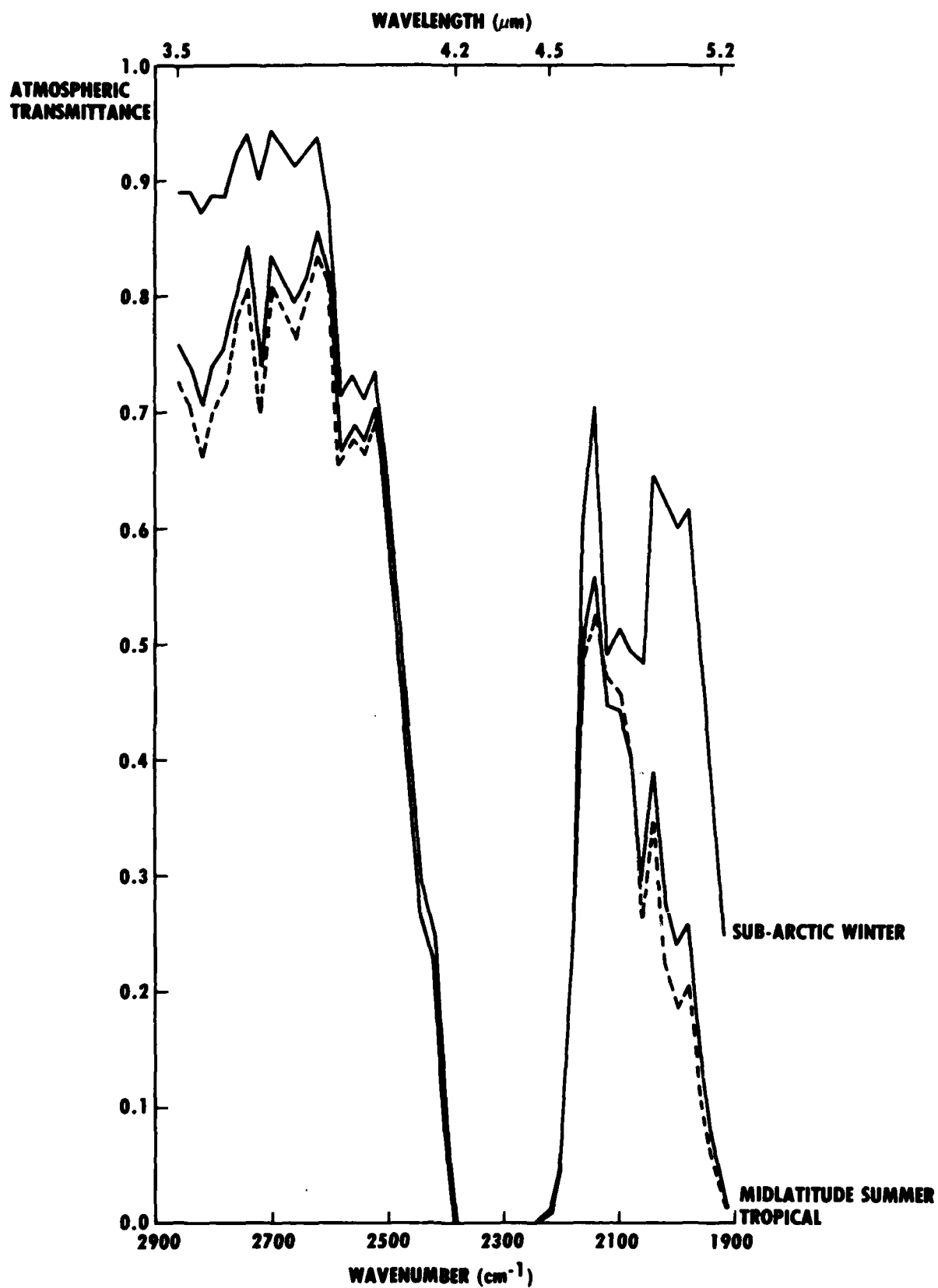
Figure 5. Atmospheric transmittance for 100 km path length between 5 and 11 km altitude for three types of atmosphere.
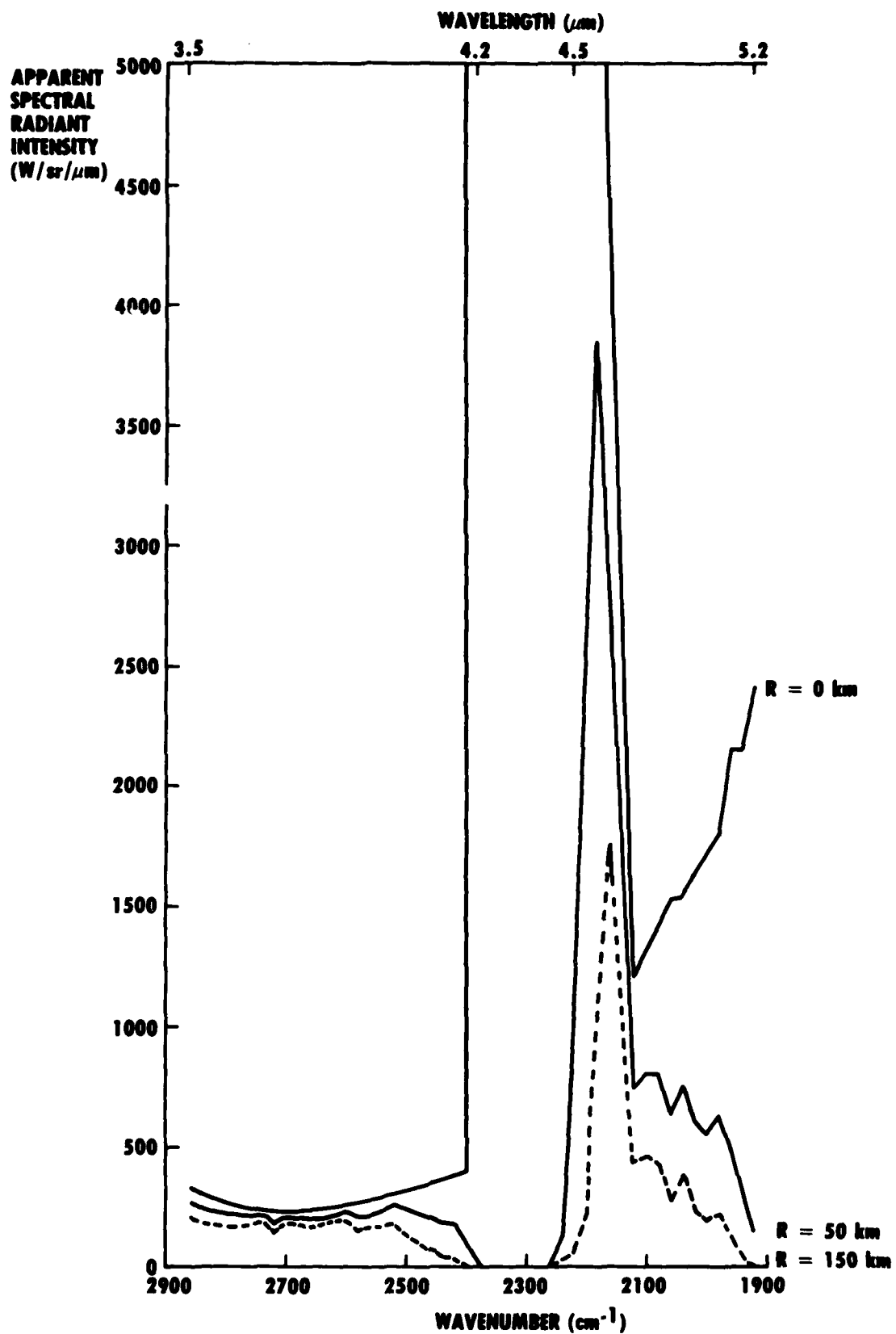
Figure 6. Apparent spectral radiant intensity of afterburning aircraft as seen at various ranges through a tropical atmosphere.

10⁻⁹

10⁻¹⁰

3.5 - 5.2 μm

4.5 - 5.2 μm

ΔH   10⁻¹¹

3.5 - 4.2 μm

10⁻¹²

SENSORS NEI IN EXAMPLE

10⁻¹³

0 10 20 30 40 50 60 70 80 90 100 110 120 130 140 150

R (km)
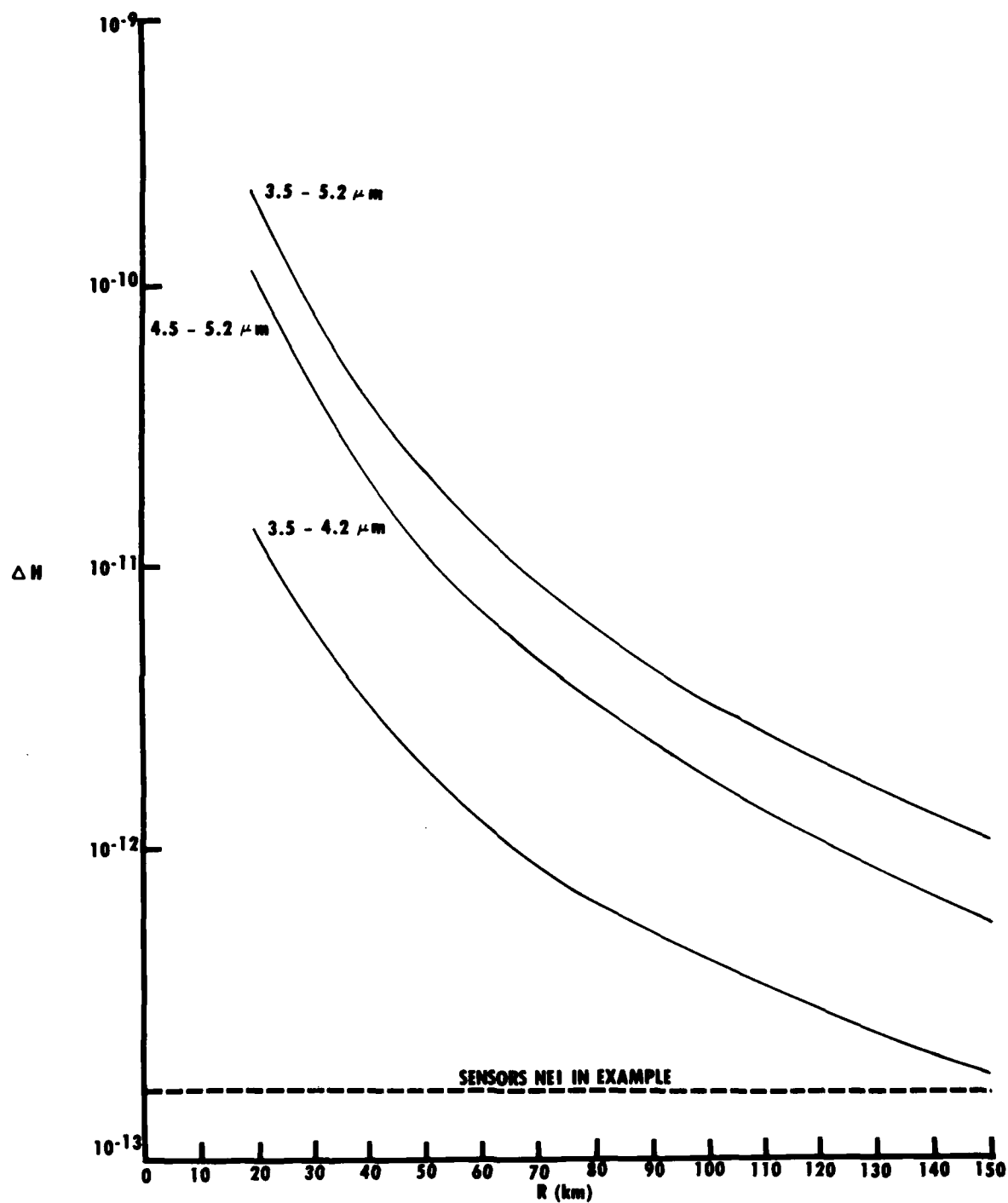
Figure 7.   Contrast irradiance versus range curves for an afterburning aircraft against a clear sky
background in various wavebands.
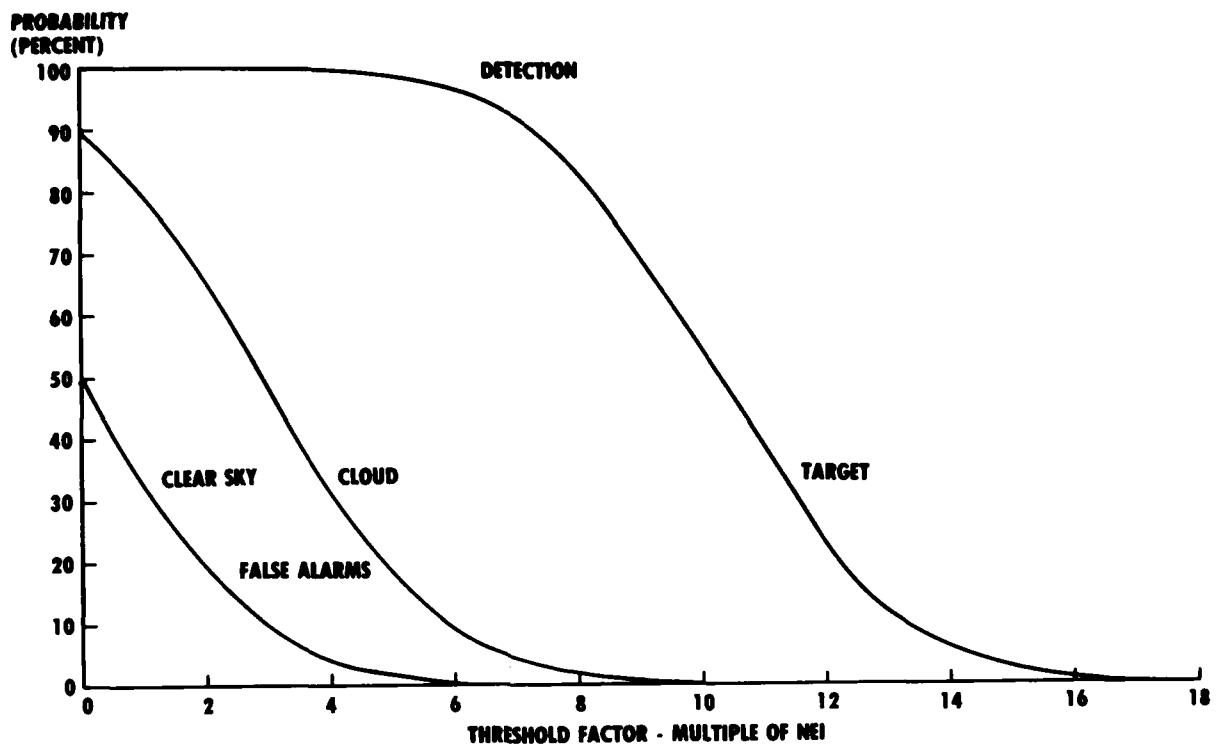
Figure 8. Probabilities of detection and false alarm as functions of threshold setting for a range of 100 km.

DETECTION ET CLASSIFICATION DE CIBLES EN IMAGERIE INFRA-ROUGE

Jean LOUCHET

Etablissement Technique Central de l'Armement
94114 ARCUEIL CEDEX , FRANCE

Cet exposé est consacré à la définition d'une chaîne de traitement automatique des images infra-rouges thermiques pour la détection et la classification des cibles terrestres. Dans ces images, la non-uniformité du fond et la multiplicité éventuelle des cibles rendent nécessaires des traitements relativement complexes, aboutissant à l'extraction des silhouettes des cibles.

Les méthodes de classification des cibles sont enfin abordées.

INTRODUCTION

Le Laboratoire de Traitements d'Images a entrepris une étude consistant à mettre au point des algorithmes de traitement automatique d'images infrarouges destinés à la détection des cibles terrestres.

Les traitements sont effectués sur des images fournies par des capteurs fonctionnant à des longueurs d'onde voisines de 4 µ et 10 µ.

Ces images diffèrent des images dans les longueurs d'onde visibles, sous deux aspects :

- d'une part, leur qualité généralement inférieure en résolution et en rapport signal à bruit, dûe aux limites physiques et technologiques des capteurs ;

- d'autre part, une qualité essentielle pour les applications militaires, qui provient de ce que le rayonnement propre des objets aux températures ambiantes est, aux longueurs d'onde voisines de 10 µ, généralement supérieur au rayonnement réfléchi : ce qui fait que ces images contiennent essentiellement une information sur la température des différents points de la scène observée.

L'étude de ces algorithmes s'effectue au moyen du système interactif de traitement du Laboratoire de Traitement d'Images de l'ETCA, couplé à l'ordinateur CIMSA-MITRA 15 du Centre Informatique de l'ETCA.

Les procédés de traitement présentés sont adaptés aux images terrestres dans lesquelles la non-uniformité du fond et la multiplicité éventuelle des cibles rend nécessaire un traitement relativement complexe. L'application de ces méthodes à la base de données " ALABAMA*" est donnée en fin de rapport (paragraphe 4.3).

Le traitement comporte trois parties :

- Les prétraitements, destinés à diminuer l'effet de certains défauts qui affectent l'image originale et qui sont dus aux imperfections du capteur (chapitre 1) ;

- La détection des cibles, consistant à classer chacun des points de l'image parmi les deux catégories : fond ou cible (chapitre 2) ; cette étape du traitement laisse généralement subsister un grand nombre de fausses alarmes qu'il s'agit de distinguer des cibles réelles ;

- l'élimination des fausses alarmes , destinée à affiner le traitement de l'étape précédente (chapitre 3).

L'enchaînement de ces trois parties,l'application à la base de données " ALABAMA " et le bilan sont donnés au chapitre 4.

---

* Il s'agit d'un ensemble de 43 images infrarouges thermiques servant de référence à plusieurs pays de l'OTAN afin de pouvoir comparer leurs méthodes de traitement d'images.

I - PRETRAITEMENTS

1.1 - Niveaux moyens des lignes (figure 2)

On constate visuellement un double défaut affectant chacune des images sur l'ensemble de leur surface : un défaut moyen local de niveau sur les lignes, avec une période de 10 lignes, et sur chaque ligne un défaut de niveau sur les points, avec une période de 3 points si l'on se restreint à une petite partie de la ligne (voir figure 2).

Le défaut de niveau local sur les lignes est un défaut courant sur les images infrarouges ; il est dû au fait que l'image est obtenue par un déplacement d'une barrette (ou d'une matrice) de photodiodes et que les dynamiques des différents détecteurs ne sont pas identiques.

Une méthode simple a été testée pour corriger ce défaut par addition d'une constante calculée pour chacune des lignes de l'image
L'hypothèse selon laquelle l'erreur sur les lignes est purement additive, s'est montrée trop restrictive : en effet, on constate que, si le résultat est bon lorsqu'on travaille sur une fenêtre de petite taille, dans le cas contraire, le décalage moyen de niveau d'une ligne par rapport à ses voisines ne reste pas constant le long de cette ligne.

Nous avons tenté de résoudre ce problème, d'abord par deux méthodes fondées sur le calcul d'un ensemble de fonctions de transcodage : on a constaté encore une amélioration par rapport à la méthode précédente, mais des défauts, visuellement de même nature, subsistent.

La méthode qui s'est finalement montrée la mieux adaptée est la suivante .

Soit b_0 le point courant, et soient les trois portions de lignes consécutives de l'image, contenant les points voisins de b_0 :

$$a_{-9} \; a_{-8} \cdots a_{-3} \; a_{-2} \; a_{-1} \; a_0 \; a_1 \; a_2 \; a_3 \; a_4 \; a_5 \; a_6 \; a_7 \; a_8 \; a_9$$
$$b_{-9} \; b_{-8} \cdots b_{-2} \; b_{-1} \; b_0 \; b_1 \; b_2 \cdots b_9$$
$$c_{-9} \; c_{-8} \cdots c_{-2} \; c_{-1} \; c_0 \; c_1 \; c_2 \cdots c_9$$

On calcule les moyennes :

$$A = \frac{1}{19} \sum_{-9}^{+9} a_i$$

$$B = \frac{1}{19} \sum b_i$$

$$C = \frac{1}{19} \sum c_i$$

calculées avec les termes a_i, b_i, c_i de l'image originale.

La nouvelle valeur calculée de b_0 est :

$$b'_0 = b_0 + \left(\frac{A+C}{2} - B \right)$$

La correction du défaut est donc adaptative le long d'une ligne.

Cette méthode donne de bons résultats ; elle serait probablement encore meilleure si elle prenait en compte 5 lignes au lieu de 3, avec une pondération appropriée, mais le temps de calcul sur MITRA 15 se ressentirait du nombre d'entrées-sorties. Elle ne tient pas compte de la "périodicité" de 10 et peut donc être utilisée pour toutes sortes d'images présentant des défauts de lignes.

L'optimisation pour un capteur du nombre des points sur lesquels se fait la pondération (19 dans l'exemple ci-dessus), a été déterminée empiriquement.

1..2 - Bruit ligne

Le bruit ligne présente une périodicité apparente de 3 points.

Là encore, cette "période" n'existe qu'à petite échelle et la méthode qui consiste à calculer trois fonctions de transcodage (selon le résidu modulo 3 de l'abscisse du point) échoue complètement.

Ce défaut n'est donc pas du tout de même nature que le défaut de niveaux moyens des lignes décrit précédemment.

Nous avons étudié les causes possibles de ce défaut et formé l'hypothèse suivante, qui paraît fort vraisemblable, selon laquelle le capteur, qui fournit des images de résolution médiocre, est muni d'un dispositif de traitement du signal destiné à améliorer cette résolution, et qui, en guise d'amélioration, introduit dans l'image le bruit que nous cherchons à corriger. La solution à notre problème consistera donc à reconstituer l'image avant l'"amélioration".

Nous avons donc simplement programmé un "retour à l'image originale", (qui, si elle est moins nette, est aussi beaucoup moins bruitée), par un flouage très modéré des lignes.
b_0 étant le point courant, b_{-1} et b_1 ses voisins :

$$b_{-1} \quad b_0 \quad b_1$$

la nouvelle valeur calculée pour b_0 est :

$$b'_0 = \frac{2b_0 + b_{-1} + b_1}{4}$$

Ceci revient non pas exactement à un retour à l'image originale, mais à un déflouage moins brutal que celui qui est fait dans le capteur.

2 - DETECTION DES CIBLES

2.1 - Introduction

Le but que nous recherchons est d'obtenir, à partir de l'image originale, une image binaire en tout-ou-rien, représentant au mieux la silhouette des cibles.

L'expérience montre qu'un seuillage uniforme effectué sur l'image originale donne des résultats médiocres, et que, même lorsque le niveau de seuillage est déterminé "à la main" par un opérateur, il n'est souvent pas possible d'avoir simultanément une extraction correcte de silhouette et un taux raisonnable de fausses alarmes (figure 4).

Un remède consiste donc à appliquer à l'image un seuillage dont le niveau, au lieu d'être constant, s'adapte localement au niveau moyen de l'image. On réduira ainsi le taux de fausses alarmes et on améliorera le compromis entre les niveaux de seuillage idéaux des différentes régions de l'image.

2.2.1 - Mode opératoire

En pratique, nous avons choisi une méthode équivalente à celle-ci, qui consiste à soustraire à l'image de départ, sa composante à basse fréquence spatiale, puis à effectuer un seuillage conventionnel.

2.2.2 - Choix de la fréquence de coupure

Le choix de la fréquence de coupure (c'est-à-dire de la largeur du flou) nécessite que l'opérateur fournisse à la machine une valeur, qui sera par exemple le diamètre de la cible, exprimé en nombre de points de digitalisation. On constate que cette valeur n'a besoin d'être connue qu'avec une précision d'un facteur 2 ou 3 sans que cela n'affecte sensiblement les résultats. On peut donc estimer que l'"aide extérieure à l'algorithme" n'est pas abusive, d'autant plus que la présence simultanée d'un télémètre est probable sur un système utilisant des algorithmes complexes comme ceux qui font l'objet de ce rapport.

⌐Pour cela, on exécute la procédure suivante :

```
                                    ┌─────────────────┐
                                    │ Image de départ │
                                    └─────────────────┘
              ┌──────┐
              │ Flou │
              └──────┘
        ┌──────────────────────┐         ╭──────────────╮
        │     image floue      │────────▶│ Soustraction │
        │ (= image des basses  │         │   d'images   │
        │ fréquences spatiales)│         ╰──────────────╯
        └──────────────────────┘                │
                                                 ▼
                                    ┌─────────────────────┐
                                    │  image de départ    │
                                    │  - image floue      │
                                    │ (= image filtrée par un
                                    │    filtre spatial   │
                                    │     coupe-bas)      │
                                    └─────────────────────┘
```
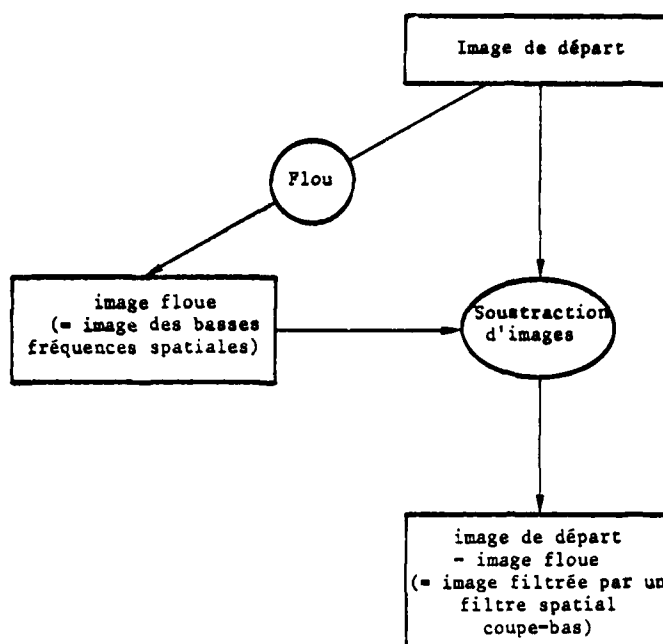
2.3 - Seuillage automatique

2.3.1- Méthode de calcul du seuil de binarisation

La méthode de calcul du seuil que nous avons mise au point est la suivante :

- on suppose connaître a priori l'ordre de grandeur ℓ de la taille * des cibles (cette information équivaut à une information sur la distance de prise de vue, connaissant la distance focale de l'optique utilisée et le tramage de la caméra infrarouge utilisée : voir paragraphe 2.3.3)

- on partitionne l'image en N carrés dont le côté est égal à ℓ,

- à chaque carré, on associe le niveau de gris le plus fort de ce carré (c'est-à-dire le niveau de gris du point le plus brillant).

On obtient ainsi un tableau T de N nombres compris entre 0 et 255.

La méthode la plus simple consiste alors à choisir pour seuil le plus petit élément de T (méthode du "seuillage par minimax") . Cette méthode fonctionne bien pour des images très simples, comportant une ou plusieurs cibles chaudes sur un fond uniforme : en effet, si nous appelons "point d'alarme" un point de niveau supérieur au seuil, le seuil choisi par minimax sera le seuil le plus bas tel qu'il existe au moins un carré ne contenant aucun point d'alarme.

Mais dès que le fond est non-uniforme (par exemple, sol foncé et ciel clair) cette méthode échouera car, pour peu qu'il existe un carré entièrement contenu dans la région de l'image correspondant au sol, le ciel sera détecté comme cible.

Nous avons donc choisi une amélioration de la méthode, fondée sur le raisonnement suivant :

- si comme supposé, les carrés ont environ la même taille que les cibles, et s'il y a P cibles, il y aura généralement 4P carrés contenant des points de cibles,

- si K est la probabilité pour que, dans un carré, il y ait une fausse alarme, il y aura environ KN carrés contenant des points de fausses alarmes. (K est évidemment très inférieur à 1),

- il y a donc, si l'image est convenablement seuillée, 4P + KN carrés contenant des points supérieurs au seuil.

Le seuil sera donc choisi à la valeur de l'élément du tableau T tel qu'il y ait, dans le tableau, N' = KN + 4P éléments supérieurs à lui (on supposera N'< N).

---

*diamètre apparent, exprimé en nombre de points de digitalisation.

Les paramètres P et K sont à estimer à priori et une fois pour toutes en fonction des images disponibles, de façon à obtenir un bon compromis entre les différentes images, et de façon que, dans une image donnée, le niveau de seuillage soit affecté le moins possible par la taille de la fenêtre de travail sur laquelle le seuil est calculé.

Les essais effectués sur de nombreuses images provenant de bases de données différentes ont montré que les paramètres :

$$K = \frac{1}{4} \, , \; P = 7$$

constituent un bon compromis.

Dans la pratique, cette formule donne pour graphe de la fonction : $N \rightarrow N'$, une droite, et n'est valable que pour les grandes valeurs de N (sinon on risque d'obtenir $N' > N$).

Afin de l'étendre à toutes les valeurs de N, nous avons introduit une saturation par la formule :

$$N' = \min \; \left(\frac{N}{4} + \frac{30N}{N + 25} \, , \; N-1 \right)$$

$$\text{soit } N-N' = \max \; \left(\frac{3N}{4} - \frac{30N}{N + 25}, 1 \right)$$

Les 100 premières valeurs de cette fonction sont représentées à la figure 10.

Le calcul du seuil est assuré par le sous-programme THRESH (voir annexe I).

2.3.2 - Seuillage

Le seuillage ayant été calculé par la méthode du paragraphe 2.4.1, on procède au seuillage de l'image, consistant à calculer à partir de l'image de départ, une image à deux niveaux de gris :

- si un point de l'image de départ a un niveau de gris supérieur au seuil, on donne la valeur 0 au point homologue de l'image seuillée ,

- si un point de l'image de départ a un niveau de gris inférieur au seuil, on donne la valeur 255 au point homologue de l'image seuillée,

On obtient donc - et c'est principalement ici qu'intervient le fait que dans les images infrarouges thermiques les niveaux de gris des points des cibles sont plus élevés en moyenne que ceux des points du fond - une image à deux niveaux (0 et 255) dans laquelle :

- les points des cibles sont à la valeur 0,

- les points du fond sont à la valeur 255.

Ceci constitue le cas idéal ; dans la pratique beaucoup d'imperfections s'introduisent, qui affectent la qualité des silhouettes détectées (trous, cibles en morceaux), ou, beaucoup plus nombreuses, qui correspondent à des fausses alarmes.

Le seuillage doit donc être affiné par une série de traitements destinés à remédier aux défauts précédents.

Ces traitements font l'objet du chapitre 3.

3 - ELIMINATION DES FAUSSES ALARMES

Le seuil ayant été calculé pour chacune des images obtenues par les méthodes des paragraphes 2.2 et 2.3, une image binaire a été finalement obtenue. L'élimination des fausses alarmes, qui est la première étape de l'identification, nécissite non plus un traitement ligne par ligne, mais cible par cible. Il convient donc de numéroter ces cibles, qui sont définies comme les ensembles connexes de points d'alarme.

3.1 - Elimination des fausses alarmes ponctuelles et rebouchage

L'image binarisée par le seuillage se présente maintenant sous la forme suivante :

- point dont la valeur dans l'ancienne image était inférieure au seuil : 255,

- point dont la valeur dans l'ancienne image était supérieure au seuil : 0.

Autrement dit, les points du fond sont au niveau 255, et les points de cible au niveau 0 (aux fausses alarmes et aux défauts de détection près).

L'image est donc prête à subir le programme de numérotation des composantes connexes (voir paragraphe 3.2). Mais, à ce stade, l'image contient généralement un grand nombre de fausses alarmes ponctuelles, qui seraient normalement éliminées aux étapes ultérieures, mais qui ralentissent considérablement l'étape suivante du traitement.

Nous procédons donc à une étape de <u>filtrage</u> destinée à <u>simplifier l'image</u> en éliminant les alarmes ponctuelles ou quasi-ponctuelles. Ce filtrage est fondé sur le principe suivant : étant donnée la configuration de points de l'image :

$$a\ b\ c\ d\ e$$
$$f\ g\ h\ i\ j$$
$$k\ l\ m\ n\ p$$

où (l'image étant binaire) tous les points valent 0 ou 255, on compare la somme =

$$S = a+b+c+d+e+f+g+i+j+k+l+m+n+p$$

(le point central h étant excepté)

aux nombres : 255xM et 255xM

où M et N sont les paramètres du filtre $(2 \leqslant M < N \leqslant 12)$

Ce filtrage a l'avantage de déformer peu les grandes cibles et d'assurer cependant une élimination efficace et un bon rebouchage des trous subsistant dans les cibles, à condition de bien choisir les paramètres (voir figure 11).

3. - <u>Critères d'élimination des fausses alarmes</u>

A ce stade du traitement, l'image est prête pour le calcul des paramètres qui concernent chacune des cibles détectées. On a en effet obtenu l'image segmentée qui contient la numérotation des cibles, et on a évidemment pris soin de conserver,non pas exactement l'image originale, mais l'image qui résulte de sa restauration (chapitre 1) , toute l'information que nous avons perdue dans la compression d'information que constitue la détection des silhouettes des cibles, est en fait encore disponible.

Nous allons donc pouvoir maintenant faire intervenir certains critères dont, jusqu'ici, nous n'avons pas pu tenir compte.

3.1 - <u>Critères de forme</u>

Ces critères font intervenir uniquement la forme du contour de l'objet détecté.

Les cas les plus fréquents de fausses alarmes dont l'élimination relève simplement des critères de contours sont :

a) - Objet entièrement contenu dans deux lignes d'image ou deux colonnes d'image consécutives,

b) - Objet peu dense ou très ellongé,

c) - Objet occupant un très grand nombre de lignes ou de colonnes (plus de 4 fois la taille a-priori).

3.2 - <u>Critères de frontière</u>

Après l'élimination précédente, qui permet de supprimer les fausses alarmes "évidentes", restent le plus douvent des fausses alarmes dont la forme et les dimensions ne permettent plus de les distinguer des cibles réelles. C'est là que nous pouvons faire intervenir un critère complémentaire des précédents, plus lourd mais très puissant,faisant appel aux données de l'image originale. Le principe en est le suivant : une fausse alarme est un objet détecté par erreur par le seuillage, et dont les bords n'ont aucune signification sémantique, alors que le contour détecté d'une cible est constitué de points du bord de la cible, donc, de points où le contraste local est fort dans l'image originale.

Dans l'image segmentée, on ne voit aucune différence entre ces deux types de contours ; si l'on remonte à l'image originale, on reconnaîtra une cible d'une fausse alarme à ce que la valeur moyenne du gradient sur son contour est très forte.

On calcule donc une image de gradients de l'image originale (on choisit un "gradient de frontière" tel qu'il est défini dans [I] , annexe II), puis pour chaque objet, les paramètres :

- IPERI, périmètre de l'objet

- SGRAD, intégrale du gradient le long du contour de l'objet

- ICONT = SGRAD/IPERI = moyenne du gradient sur le contour de l'objet.

L'algorithme peut être résumé par le schéma suivant :

```
                    ┌─────────────────────────────────────┐
                    │ DECADRAGE ET CREATION D'UNE MARGE     │
                    └─────────────────────────────────────┘
                                      │
                    ┌─────────────────────────────────────┐
   Restauration de  │ CORRECTION DES DEFAUTS DE LIGNE      │
   l'image bruitée  └─────────────────────────────────────┘
                                      │
                    ┌─────────────────────────────────────┐
                    │ REDUCTION DU BRUIT LIGNE             │
                    └─────────────────────────────────────┘
                                      │
                    ┌─────────────────────────────────────┐
                    │         FILTRAGE SPATIAL            │
                    │           COUPE-BAS                 │
                    │      (création d'une image          │
   Détection        │      floue et soustraction)         │
      des           └─────────────────────────────────────┘
    cibles                            │
                    ┌─────────────────────────────────────┐
                    │ CALCUL DU SEUIL ET SEUILLAGE        │
                    └─────────────────────────────────────┘

                    ┌─────────────────────────────────────┐
                    │ FILTRAGE LOGIQUE DE L'IMAGE BINAIRE  │
   Mise en          └─────────────────────────────────────┘
   forme                            │
                    ┌─────────────────────────────────────┐
                    │ NUMEROTATION ET COMPTAGE DES CIBLES  │
                    └─────────────────────────────────────┘

                 ┌──────────────────────────────────┐   ┌───────────────┐
   Elimination   │ CALCUL DES PARAMETRES CARACTERISTIQUES │ │ CALCUL DE L'IMAGE │
   des fausses   └──────────────────────────────────┘   │      DES      │
   alarmes       ┌──────────────────────────────────┐   │   GRADIENTS   │
                 │ ELIMINATION DES FAUSSES-ALARMES    │   └───────────────┘
                 └──────────────────────────────────┘

                 ┌──────────────────────────────────┐
   Identi-       │ CALCUL DES PARAMETRES D'IDENTIFICATION │
   fication      └──────────────────────────────────┘
                                │
                    ┌────────────────────┐
                    │   IDENTIFICATION    │
                    └────────────────────┘
```

Une étape préliminaire de "décadrage et création d'une marge" a été ajoutée en
début d'enchaînement afin de rendre possible le calcul du filtrage spatial passe-
bas (flou, voir paragraphe 2.3.2), sans avoir d'effet de bord (voir figure 13).

Elle consiste à entourer l'image originale d'une marge, de largeur au moins égale à
la largeur du flou à calculer, dont la valeur de chaque point est obtenue par reco-
pie du point correspondant situé sur le bord de l'image originale.

   ## Résultats

Le programme de détection des cibles et d'élimination des fausses alarmes a été
testé sur les images n° 1 à 20 de la base de données "Alabama". Les résultats sont
les suivants :

- Sur 39 cibles, 5 n'ont pas été détectées et 2 fausses alarmes se sont introdui-
  tes, ce qui correspond à :

                    taux de détection : 87 %

                    taux de fausses alarmes : 5 %

  à condition de considérer comme correcte la détection des cibles qui se recou-
  vrent.

- Cette série ne comporte que trois images dans la bande 4 microns, ce qui est
  peu significatif, mais on remarque que, sur ces images, sur 6 cibles, 4 ont été
  détectées, 2 perdues et 1 fausse alarme s'est déclarée.

5 - CLASSIFICATION

Une méthode de classification par corrélation avec apprentissage est en cours
d'étude. L'algorithme d'apprentissage, décrit plus bas, consiste à déterminer
la meilleure approximation du paramètre de reconnaissance (par exemple objet de Type A =0,
objet de type B = +1)   comme combinaison linéaire des paramètres définissant
l'image à étudier. Cet algorithme a été testé sur des images simples et fonc-
tionne de façon satisfaisante ; son application étant pour des raisons matériel-
les limitée aux images ne comportant pas plus d'une trentaine de points, il est
nécessaire de comprimer l'information liée à une cible à une trentaine de para-
mètres, et ceci de façon stable par rapport aux variations dues aux changements
des conditions de prise de vue (angle de prise de vue, etc...). Nous n'avons pas
encore étudié la façon de comprimer cette information ; il semble qu'un codage
du contour constitue la meilleure voie, bien que d'autres méthodes soient envi-
sageables.

Principe de l'algorithme d'apprentissage

Une fois choisie une méthode de compression de l'information permettant d'asso-
cier n paramètres descriptifs à chaque image de cible, il convient de trouver
une méthode qui permette de classifier la cible décrite par ces n paramètres que
l'on nomme "paramètres objectifs" L identification se fera par calcul de p pa-
ramètres (du type : $x_i = 0$ si objet du type A

$x_i = 1$ si objet du type B, etc)

Ces p paramètres se nomment "paramètres subjectifs".

Le problème consiste à exprimer les p paramètres subjectifs en fonction des n
paramètres objectifs. La solution donnée ici permet de calculer les coefficients
de l'application linéaire répondant le mieux à la question.

Pour cela, l'opérateur doit choisir une série représentative d'images d'entraîne-
ment dont il connait la vérité terrain.

Pour chacune des cibles de ces images, l'ordinateur calculera les n paramètres
objectifs selon la méthode de compression choisie, et l'opérateur fournira lui-
même les p paramètres subjectifs décrivant la vérité terrain.

Le rôle de l'algorithme d'apprentissage est de calculer, à partir de ces données
objectives et subjectives, qui seront généralement corrélées, les n x p coeffi-
cients demandés.

Algorithme d'apprentissage

Soit une famille de m objets représentés par n paramètres . A chacun de
ces objets le professeur attribue p paramètres représentant la classe à
laquelle cet objet appartient. Le problème de l'apprentissage consiste à
trouver la meilleure approximation des p paramètres " subjectifs" comme
combinaison linéaire des n paramètres objectifs : ceci revient à trouver
le corrélateur linéaire qui , appliqué à l'objet représenté par ses n
paramètres, détermine la classe de l'objet.

Soient $\left\{x_i\right\}_{i = 1, n}$ les paramètres objectifs

$\left\{x_i\right\}_{i = n + 1, n + p}$ les paramètres subjectifs

Soit M la matrice de covariance des (n + p) paramètres centrés.

Notant $\tilde{x}_i = x_i - E(x_i)$ on a :

$$M = \left[E(\tilde{x}_i \ \tilde{x}_j) \right] \quad \text{où } E \text{ est l'espérance mathématique.}$$

Classant les valeurs propres λ_i **de   dans l'ordre décroissant, on obtient
les vecteurs propres** $\vec{V}_i$ (i = 1, n + p) **de coordonnées** (V_{ij}) (j = 1, n + p)

Pour les p derniers vecteurs propres, λ_i est petit ; on peut donc écrire :

$$i = n + 1, n + p : \quad \sqrt{\lambda_i} = \sigma \left(\sum_{j=1}^{n+p} V_{ij} \; \tilde{x}_j \right) \simeq 0 \quad \text{où } \Gamma \text{ est l'écart-type.}$$

Par conséquent, si l'on écrit les p équations $(i = n + 1, n + p)$:

$$\sum_{j=1}^{n} V_{ij} \; \tilde{x}_j + \sum_{j=n+1}^{n+p} V_{ij} \; \overset{::}{x}_j = 0$$

Les solutions $\overset{::}{x}_j$ seront des approximations des $\tilde{x}_j$ $(j = n + 1 \text{ à } n + p)$

On obtient donc le système de p équations linéaires à p inconnues $x_j^{\dagger}$

$$(i = n + 1, n + p) \quad \sum_{j=n+1}^{n+p} V_{ij} \; \overset{::}{x}_j = - \sum_{j=1}^{n} V_{ij} \; \tilde{x}_j$$

dont la résolution donnera pour chaque valeur $i = n + 1, \ldots n + p$:

$$(i = n + 1, n + p) \quad \overset{::}{x}_i = \sum_{j=1}^{n} W_{ij} \; \tilde{x}_j$$

définissant le corrélateur idéal.

Appliqué à un objet S quelconque, ce corrélateur fournit les p paramètres.

$$\overset{x}{x}_i (S) = \sum_{j=1}^{n} W_{ij} \; \tilde{x}_j (S),$$

prédictions des paramètres subjectifs, en fonction des paramètres objectifs centrés $(\tilde{x}_j (S))_{j=1,n}$

N.B. Contrairement à ce que l'on observe le plus souvent, nous n'avons pas démontré l'existence d'un corrélateur optimal mais, ce qui est plus utile, nous avons donné une méthode de calcul d'un corrélateur. Ceci étant, on ne peut parler d'optimalité que dans le mesure où on ne l'applique qu'aux objets ayant servi à l'apprentissage.

La réussite de la méthode sera donc liée :
- d'une part, à la représentativité de l'échantillon choisi pour l'apprentissage;
- d'autre part, à la bonne adaptation de la compression d'information utilisée pour calculer les n paramètres objectifs. Une mesure de cette bonne adaptation sera le degré de corrélation entre paramètres objectifs et subjectifs, que l'on peut mesurer par la bonne ou mauvaise décroissance des valeurs propres.

5 - CONCLUSION

Cette partie de l'étude a permis de mettre au point un algorithme de détection
automatique des cibles et de leurs contours, les essais effectués sur la base de
données "Alabama" ont montré qu'il fonctionnait de façon satisfaisante dans la
majorité des cas, puisque l'on atteint un taux de détection de 87 % pour un taux
de fausses alarmes de 5 %.

L'étape suivante de l'étude, conduisant à la classification automatique des cibles,
est actuellement en cours.

ANNEXE - ELIMINATION DES FAUSSES ALARMES

00013

00018   00045 → Valeur maximale du gradient sur l'image = MAXGRAD (si ICONT < MAXGRAD/3, cible rejetée)
                                                                ici, MAXGRAD/3 = 15

   └─►Taille a priori des cibles (voir paragraphe 2.3.3)

```
00001 00118 00121 00215 00220 00004 00006 00024 00012 00017 00003 00000 00255
00002 00299 00303 00197 00205 00005 00009 00045 00034 00025 00004 00000 00255
00003 00307 00308 00311 00314 00002 00004 00008 00007 00011 00002 00000 00255
00004 00308 00309 00269 00272 00002 00004 00038 00006 00011 00003 00000 00255
00005 00309 00309 00277 00278 00001 00002 00002 00002 00005 00002 00000 00255
00006 00319 00320 00125 00130 00002 00006 00012 00010 00013 00002 00000 00255
00007 00332 00337 00073 00283 00006 00011 00066 00034 00034 00003 00000 00255
00008 00332 00334 00090 00097 00003 00008 00024 00017 00019 00003 00000 00255
00009 00332 00333 00098 00101 00002 00004 00008 00207 00010 00005 00000 00255
00010 00333 00342 00273 00282 00010 03010 00100 00072 00035 00027 00001 00001   - cible n° 1
00011 00334 00348 00144 00166 00015 00023 03345 00232 00065 00027 00001 00002   - cible N° 2
00012 00336 03348 00179 00196 00013 00018 00234 00188 00055 00031 00001 00003   - cible n° 3
00013 00337 00339 00250 00054 00003 00005 00215 00010 03013 00004 00000 00255
```

Colonnes (de droite à gauche) :

- Nouveaux numéros attribués
- Nature de l'objet (cible = 1', f.a = 0')
- ICONT, moyenne du gradient sur le contour
- (voir paragraphe 3.3.2) IPERI périmètre
- SURF, aire de la cible
- ISURF, aire de la fenêtre
- COMAX-COMIN + 1, nombre de colonnes
- LIMAX-LIMIN + 1, nombre de lignes
- COMAX, N° dernière colonne
- COMIN, N° 1ère colonne
- LIMAX, N° dernière ligne
- (voir paragraphe 3.3.1) LIMIN, N° 1ère ligne
- N° de l'objet détecté

REAL-TIME GREY-LEVEL HISTOGRAM MANIPULATION

L.H. Guildford
Philips Research Laboratories
Redhill, Surrey RH1 5HA,
England.

SUMMARY

This paper considers the problem of how best to match the characteristics of a video signal of high dynamic range to those of a display. In particular, a problem often encountered in surveillance systems, that of adequately depicting small, but often important contrast changes within regions of widely differing mean brightness levels. Situations can arise where any change in overall brightness or contrast needed to reveal a particular object, if its presence were even known, would exclude details from other important areas of the scene. Details necessary for identification or even detection of the object mostly exist in areas of low contrast. The problem is to display and usefully observe a wide dynamic range of information by means of a display system whose operative grey-scale range is but a fraction of that apparently necessary to present the scene. Imaging systems operating in the visible, IR, X-ray and ultrasound regions of the spectrum fall into this category.

Published literature concerning the visibility of experimental objects against differing backgrounds is reviewed and interpreted in a manner which allows an estimate to be made of the number of grey-levels required to represent objects of varying sizes adequately on display.

A solution to the problem is proposed, which involves the application of real-time histogram modification techniques to a selected sub-area (Keyhole) within the image, combined with overall 2-D edge enhancement. Our ACE processor, which has allowed us to implement this algorithm and investigate the subjective effects of real-time operation is described and some pictorial results are presented. It has been found that this operator gives promising results especially where the next result is a local adaptive contrast stretch.

1. INTRODUCTION

Effective surveillance of a scene and quick identification of an object of interest require that the information presented to an observer has to be carefully matched to the dynamic range of the display used and has to be enhanced so that it may be rapidly assimilated by eye. Consideration has to be given not only to the limited ability of the eye to identify vehicles and other objects described by less than 13 lines (Scott, F, 1970, Hollanda, 1970), but also to the inability of a display to present sufficient perceptible grey-levels to cover the range of detectable signal-levels produced by many imaging systems.

In this paper we will be briefly considering the interaction of these various effects. A description will then be given of an experimental system which has enabled us to investigate both contrast modification and edge enhancement techniques applied individually or in parallel to real-time video signals. Experimental results are then presented and discussed.

In general we have restricted our activities to real-time processing. That is to say that our processing operators either work at the scanned pixel rates or else have completed their modifications within one frame period. Other constraints, such as the need to consider portability, low power and cost effectiveness have also influenced our work. As a result we have concentrated our attentions on families of local area processing operators.

As stated above, our main objective has been to bring to the observer's attention the information in the scene that was of immediate interest to him. In general, such information is characterised by low amplitude, high spatial frequency signals, often situated in areas of poor local contrast.

2. THE RELATIONSHIP BETWEEN OBJECT SIZE AND GREY-SCALE VISIBILITY

The problem of how best to display an image in order that the minimum of available information is lost is a complex one, involving consideration of both the characteristics of the display and of the human visual chain. Figure 1 shows an image that exemplifies many of the problems. Overall, the image has a wide dynamic range. However within this wide range, the objects of interest have only small contrast differences with respect to their local background. It is therefore impossible to display adequately both these small, but important local variations and the complete image. This type of dilemma is quite common - Figure 6 shows some more images which suffer from similar problems. In the region containing the hidden tank there is little detail available. The contour image, which contains only edge information, does not indicate its presence until the gain has been increased by a factor of 16. Such high contrast gains are incompatible with the satisfactory presentation of the complete picture.

One of the root causes of this problem is the mismatch between the dynamic range of the video signals and that of normal CRT displays. The chart in Figure 2 provides a simple comparison - in terms of detectable, noise limited, signal levels - between two hypothetical IR sensors, normalised to discriminate $0.1°C$, and the output from a high quality TV imager. To do full justice to the incoming information a display would have to be capable of reproducing in excess of 300 discernible grey-levels for the TV and well in excess of 400 levels for the IR systems, to do full justice to the incoming information.

However, this is too simplistic a view, since it has been found that the range of 'discernible grey-levels' observable on a display of any description is also dependent upon the angle that the objects of interest subtend at the eye.

When one considers the theoretical dynamic range of a display, one must bear in mind not only the above effect, but also the ambient lighting, background luminance, screen reflectance, flare and luminance range of the display. Figure 3 shows a series of curves which emphasise the effect of object subtense on the 'Just perceptible relative intensity' over a range of luminance values. Note that the lower luminance threshold, the 'luminance of subjective black', is itself a function of object subtense. These curves have been compiled from data and information given in reference books and reports which cover the subjective effects of various aspects of displays (Lowry, E.M., 1951, McIlwain, K.,1956, Robinson, R.N., 1972). The information contained in Figure 3 has been transcribed in Figure 4 in order to relate the number of discernible grey-levels to the subtense of the object. The curve relates to an ambient illumination of 16 lm.m^{-2} and takes into account the variation in the luminance of subjective black and the background luminance of the CRT (reflectance + flare). All the photometric measurements were undertaken using controlled illumination and surfaces of known size and reflectivity. Only the background luminance measurements involved the use of a CRT. These measurements included optical and electron flare and the reflectance of the ambient illumination (16 lm.m^{-2}) at the faceplate. However effects due to line structure and noise in the luminance signal were not investigated by these authors. The latter effect would tend to smooth out the difference between the original set of 'Just perceptible relative intensity' readings, therefore the value of $\Delta B/B$ would have to increase to make the change perceptible with noise superimposed, thus further reducing the number of discernible grey-levels.

These results, which, as we have seen, have been extrapolated from static measurements made under controlled conditions, give an arguably high estimate for the observable dynamic range of grey-levels on a display. However, they appear to agree broadly (i.e. within a factor of 2) with our own subjective estimates.

As we have already seen, much of the important information in a scene consists of small local variations superimposed on areas of often widely differing luminance. Our approach to the problem of displaying such images is to use statistical information from a given area in a scene to enable us to calculate the video transfer function which best matches it to the full dynamic range of the display. The algorithm used is based upon grey-level histogram modification (Rosenfeld, A., 1976, Hummel, R., 1977, Ketcham, D.J., 1976, Kruger, R.P., 1971). We have also found that the results of this adaptive contrast modification technique can be further improved by simultaneous 2-dimensional edge enhancement (Rosenfeld, A., 1976, Levi, L., 1976, Schreiber, W.F., 1970). In our system these two operations are performed in parallel on the separated edge and brightness signals, which are then recombined after processing to give the final output signal.

3.   EXPERIMENTAL HARDWARE

In order that the subjective effects of our proposed algorithms can be investigated in real-time, we have constructed a microprocessor controlled real-time image processor, which we call ACE (Adaptive Contrast and Edge processor), which is capable of operating on 8-bit digitized video signals at up to 15 mega pixels.s^{-1}, the more detailed block diagram is given in Figure 5.

Figure 6 however shows a much simplified block diagram of the ACE system, which is quite flexible and allows us to perform, in parallel, processing operations on both brightness and edge information in a scene.

These operations include:

Linear contrast stretch)	
Gamma correction)	Brightness processing
Thresholding)	
Histogram equalisation/modification)	
2-dimensional edge enhancement)	
Adaptive spatial filtering)	Edge processing
Adaptive noise thresholding)	

Operator interaction is via a keyboard and X-Y oscilloscope display, which is used to present plots of system transfer functions, input or output histograms and other relevant information.

High through-put and adaptivity at both pixel and frame rates are achieved by the use of parallel RAM look-up tables as the processing elements in each of the 6 main signal paths - brightness, point, horizontal, vertical and 2 diagonal edge signals, (Figure 5). Selection between the 4 parallel RAMs in each of the signal paths can be made at the pixel rate, providing limited real-time adaptivity, while the RAMs themselves are loaded by the system microprocessor, thus giving further adaptivity but in this case at frame rate. Other units ancillary to the basic processing elements are the input matrix store and point/edge generator which furnish the 6 input signals mentioned above, the path control board, which translates microprocessor mode commands into various path selection strategies for the processed signals at the pixel rate and finally, a histogram counter, which provides the necessary statistical information concerning the intensity distribution of the input brightness signal (i.e. the mean of a 3x3 pixel block surrounding the current processed

point). At the output of the system, the processed line or point and brightness signals are recombinded to give the complete output signal.

External to the ACE processor, other units are required to support it. These include an ADC and DAC, a video sync stripper, a clock generator and a Keyhole generator. The use of an accurate sync stripper, black-level d.c. clamp and clock pulse generator is important. Sync signals within the processor will only degrade its performance, using up unnecessary bits in the digitized pixel word. A black-level d.c. clamp samples the back porch of the video waveform and maintains a known d.c. reference throughout the line scan. Other points on the waveform may be sampled for reference purpose if required. The clock generator is asychronous and is capable of locking to a sync pulse on a line by line basis, so that the vertical justification of the pixel samples is maintained to within a fraction of a pixel period. Flywheel synchronisation circuits are of little use when trying to synchronise to the non-broadcast standard sync signals provided by imagers using opto-mechanical scanning systems, cheap TV cameras and video tape recorders. The Keyhole generator provides flags which delineate a variable size keyhole, which may be positioned anywhere within the frame.

It is now possible to see how a complicated algorithm such as histogram modification may be implemented on this system. At the end of a frame period, the microprocessor accesses the histogram counter, which contains a count of the number of pixels in each of the 256 input levels within the current Keyhole area. This information is used to determine the intensity transfer function that will give the best approximation to the desired output histogram. This transfer function is then loaded into one of the brightness look-up table RAMs, and the path select logic set so that this RAM is selected during the Keyhole period, a linearly coded RAM being selected when outside the Keyhole.

It is now clear why we state that edge enhancement is performed in parallel with histogram modification, since the former algorithm requires access to the point difference signal, while the latter requires the brightness (i.e. mean) signal. In the ACE system these signals are indeed separate and are modified in parallel before recombining.

4.   EXPERIMENTAL RESULTS

The pictures in this article have been derived from TV cameras or slide scanners viewing either a diorama or transparencies. In the case of the fish head (a halibut) an x-ray transparency was used. This x-ray was chosen for its wide range of grey-scales and fine detail which serve to illustrate the points we wish to emphasise.

Some idea of the subjective effect of grey-level histogram equalisation may be obtained from the rather old picture of a 256 step grey-level wedge shown in Figure 7 (Guildford, L.H., 1978/79). The upper histogram shows that within the Keyhole, the scene intensities span only one quarter of the total available brightness range. The lower histogram shows that after processing, the pixels within the Keyhole now span the entire intensity range from black to white, a contrast expansion of 4:1 in this instance. ACE implements a 'direct' histogram modification algorithm in that a count is made over all 256 grey-levels of the input signal instead of only the 32 output levels. This enables the microprocessor to calculate the optimal transfer function immediately without recourse to any iterative techniques. The modified image thus tracks any changes of the input scene on a frame to frame basis.

To illustrate the subjective effects of this form of processing we have chosen to use our 'fish head' high resolution x-ray slide. Figures 8 and 9 bring together a group of pictures each illustrating various aspects of the algorithm. The histogram modification seen in Figure 8 has in general produced a scene adaptive contrast stretch, within the programmed 32 output grey-levels to the display and has as a result brought out details that were hidden due to lack of local contrast. These effects may also be seen in Figures 1, 6 and 11. In the latter figures the transfer function reflects these changes and shows that the video transfer characteristic is a monotonically increasing indeterminate function that is scene dependent and not the more usual log, exponential or linear function associated with most video systems.

As the processing operator relies upon statistical information for its operation, loss of detail can result from adjacent differing pixels being forced into the same displayed grey-level. This may be detected in Figure 8B and is most certainly evident in the large area histogram modifications of Figure 10A. By adding 2-dimensional edge enhancement (the 'contour operator' in Figure 6) to the histogram modified video signals small differences in contrast which exist at the higher spatial frequencies are high-lighted as shown in Figures 8C and 10B. However care must be taken since one can run into noise problems if excessive edge emphasis is used.

Images which have been modified so as to give a non-rectangular histogram, such as rising or falling power law or a Gaussian distribution, are also of interest and could be of use in providing additional emphasis in a predetermined range of grey-levels. Figures 9A and B demonstrate the effect of rising and falling power law distributions on our standard picture. Emphasis of the lower and upper grey-level distributions may be observed in these scenes.

Figure 11 demonstrates the use of our histogram processor on an area of the fish head which exhibited a bi-modal grey-level distribution, see Figure 11A. The output histogram demonstrates that our operator is capable of dealing with this more complex situation. The resultant transfer function, though monotonic, exhibits 4 areas of compression and 3 areas of expansion as the histogram is in fact near to being tri-modal. In this instance the

final processed information is presented to the display in 64 grey-levels (6 bits), whereas all the previous output histogram modifications have expressed the displayed image in 32 levels.

5. CONCLUSIONS

We have found that one solution to the display dilemma discussed in the second section of this paper is to perform real-time histogram modification on a selected section of the displayed image, probably combined with limited edge enhancement. This operator provides both the local contrast stretching and adaptive compression required to match a wide range of video signals to the display. The use of keyholing techniques, however, limits the application of this form of processing to a military, medical, or scientific environment rather than to a consumer one; but it does have the advantage of allowing much greater adaptivity within the selected area, thereby bringing the most out of a scene.

The quality of the final image and the effectiveness of both processing operators, i.e. histogram manipulation and 2-dimensional edge enhancement is ultimately limited by the noise in the video waveform. A busy edge in the image is often quite acceptable to the viewer. The signal to noise ratio in the video signal from the scenes shown is of the order of 46 dB.

In the future we will be using our ACE processor to investigate the use of other non-linear grey-level histogram distributions whose selection could be made scene dependent. The use of a variable Keyhole enhancement area permits us to carry out detailed analysis of sub-frame areas, down to a single pixel if necessary. Such investigations are of interest when considering the application of grey-level modification on a pixel by pixel basis over the whole frame.

My colleague Dr P K Bailey has been a major contributer to this program of work and I would like to acknowledge his able support.

6. REFERENCES

DREWERY, J.O., 1976, Study of high-order vertical-aperture correction', Proc. IEE. Vol.123, No.4, April 1976, p.299-308.

GUILDFORD, L.H., 1978/79, 'The Dot Scan CCTV, a flexible system for real-time image-processing experiments', Philips Technical Review Vol.38, No.11/12, 1978/79, p.310-325.

HOLLANDA, P.A., et al, 1970, 'The information value of sampled images as a function of the number of scans per scene object and the signal to noise ratio', Photographic Science and Engineering, Vol.14, No.1, Jan/Feb 1970, p.407-412.

HUMMEL, R.A., 1975 'Histogram modification techniques', Computer Graphics and Image Processing, Vol.4, 1975, p.209-224.

KETCHAM, D.J., 1976, 'Real-time image enhancement techniques', SPIE/OSA, Vol. 74, 1976, Image Processing, p.120-125.

KRUGER, R.P., et al, 1971, 'Digital techniques for image enhancement of radiograms', Int. Jnl. of Bio-Medical Computing, Vol.2, 1971, p.215-238.

LEVI, L., 1976, 'On perfect image correction by unsharp masking', Proc. Int. Conference on Applications of Holography and Optical Data Processing', Jerusalem 1976, p.73-83.

LOWRY, E.M., 1951, 'The luminance discrimation of the human eye', Jnl. SMPTE, Vol.57, Sept. 1951, p.187-205.

McILWAIN, K., 1956, 'Principles of color television' by the Hazeltine Lab. staff Pub. John Wiley and Sons, 1956, Chap.11-3, 7-3.

ROBINSON, R.H., 1972, 'The contrast handling requirements of a colour television display', BBC Research Report No. 1972/37.

ROSENFELD, A., 1976, 'Digital picture processing', Academic Press NY, 1976, Chap.6.4. and 6.2.

SCHREIBER, W.F., 1970, 'Wirephoto quality improvement by unsharp masking', Pattern Recognition, Vol. 2, 1970, p.117-121.

SCOTT, F., et al, 1970, 'The information value of sampled images as a function of the number of scans per scene object', Photographic Science and Engineering, Vol.14, No.1, Jan/Feb. 1970, p.21-27.
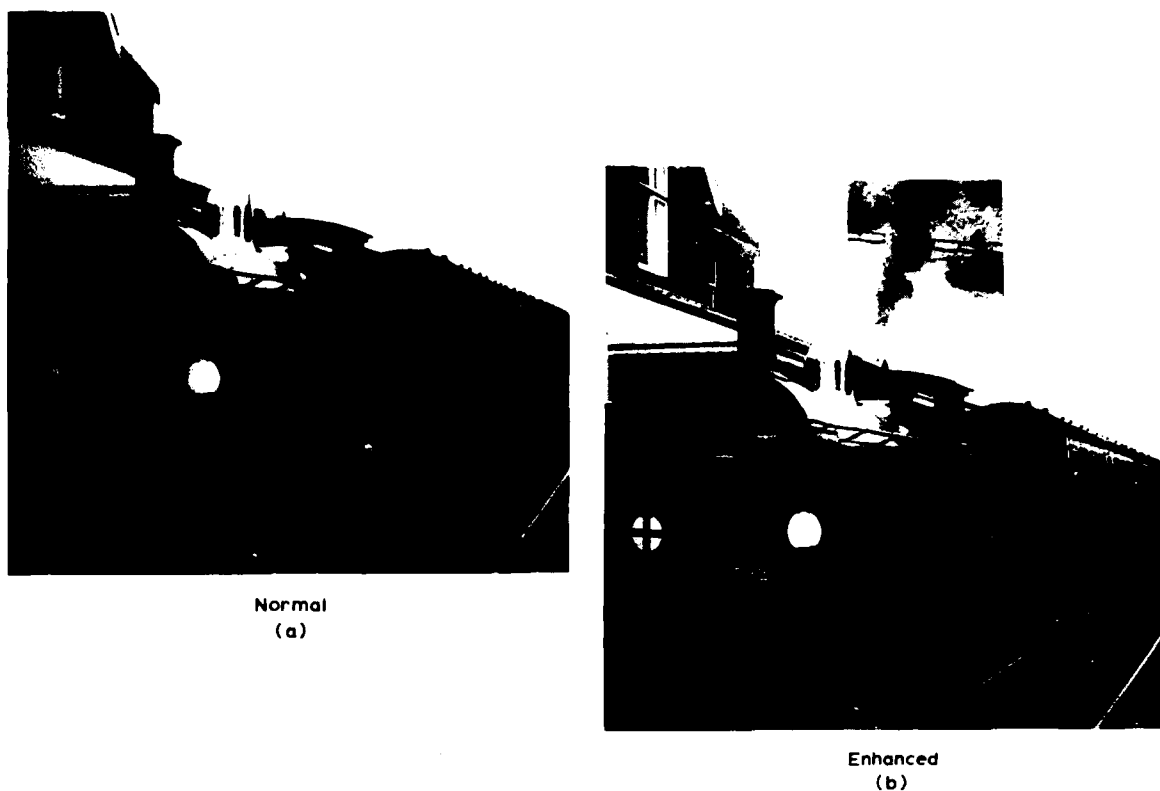
Normal
(a)

Enhanced
(b)

Fig.1 High contrast scene demonstrating loss of detail in a local area of low contrast

Luminance

T V wavelength 0.4-0.9 μm	Black		Peak white
	0		390

Object temperature   Background

IR wavelength 3-5 μm	-10	+20	+45	+70		
		370	1090	2440	0°C	
	-10	+20	+45	+70		
		270	770	1730	10°C	
	-10	+20	+45	+70		
		190	540	1270	20°C	

Object temperature

IR wavelength 8-14 μm	-10	+20	+45 +70		
		320	660	1070	0°C
	-10	+20	+45 +70		
		290	590	970	10°C
	-10	+20	+45 +70		
		260	540	880	20°C

Unweighted S/N ratios where:-

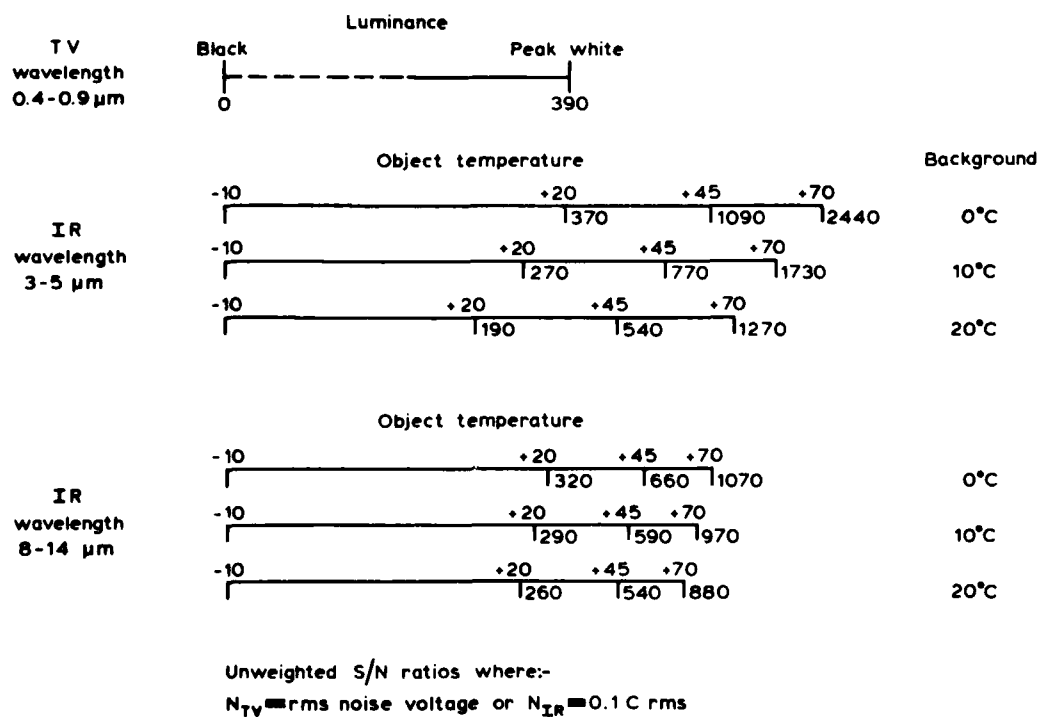N_{TV} = rms noise voltage or N_{IR} = 0.1 C rms

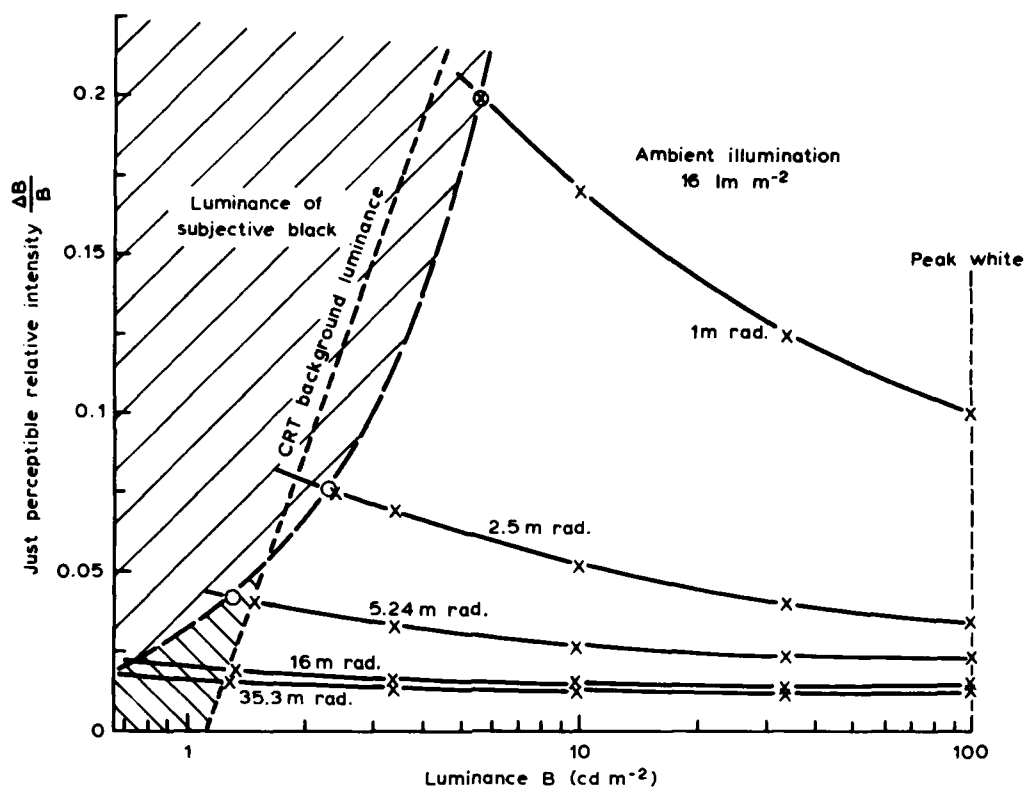Fig. 2 Detectable noise limited signal levels for TV and IR sensors

Fig. 3 Just perceptible relative intensity v. luminance, for various object subtenses - 90°/o confidence rating
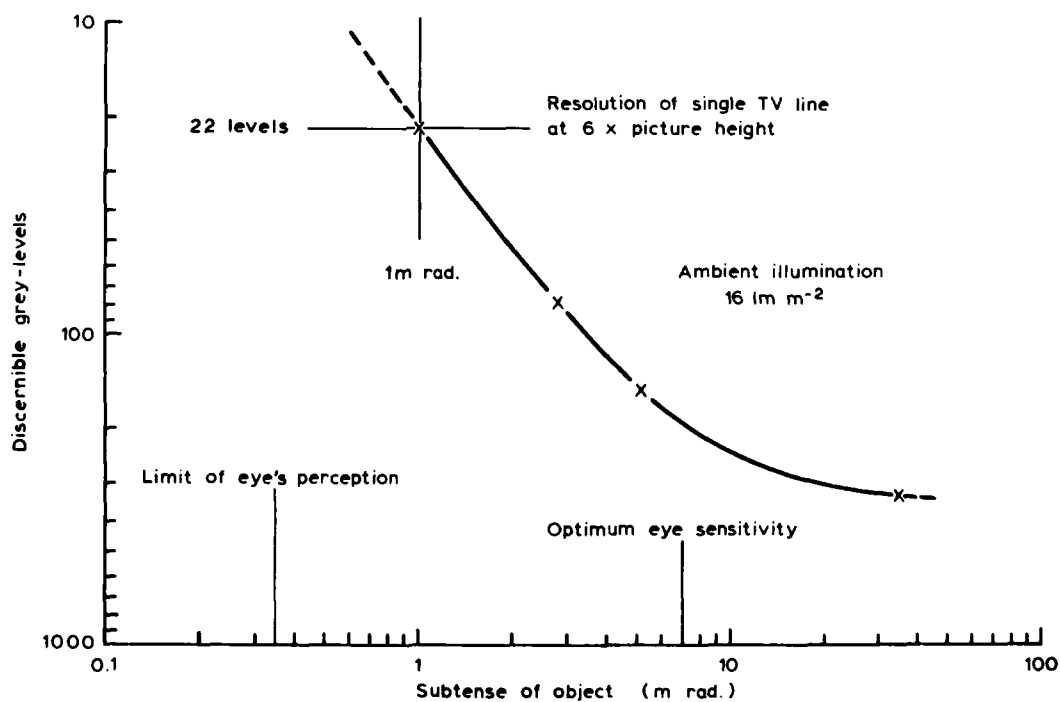

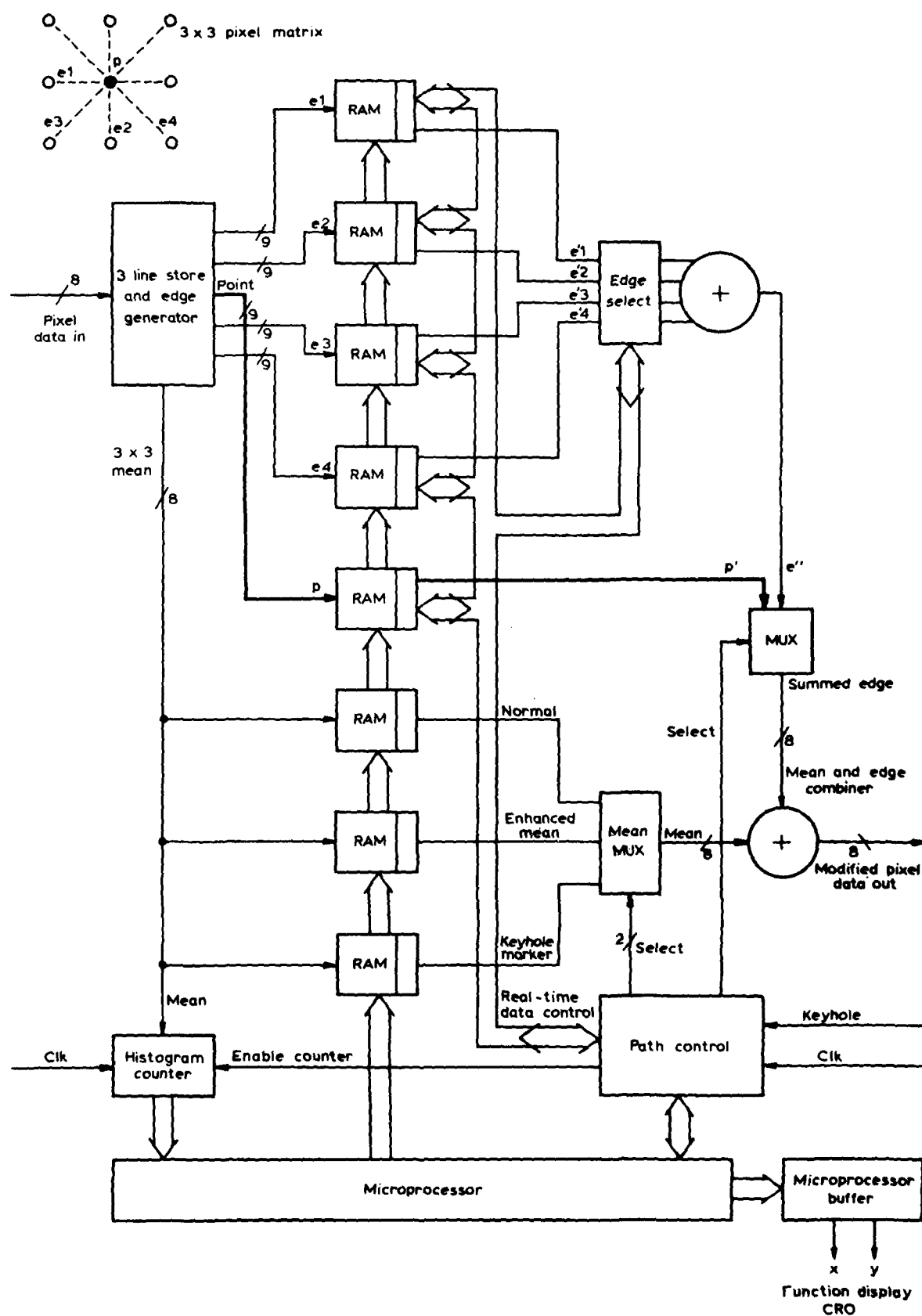Fig. 4 Discernible grey-levels v. subtense of object

Fig. 5   The microprocessor serviced, adaptive contrast and edge processor ACE

8 bit
video
data in

Input image

Input histogram

Edge and point

Edge
generator
(3 x 3 matrix)

Mean

Histogram
counter

Count
data

Contour
operator

Normal
RAM

Recode
RAM

Recode data

Microprocessor

Keyhole
control

Mean

Recoded histogram

o/p transfer function

Edge

+

Edge image

Enhanced image

8 bit enhanced
video information
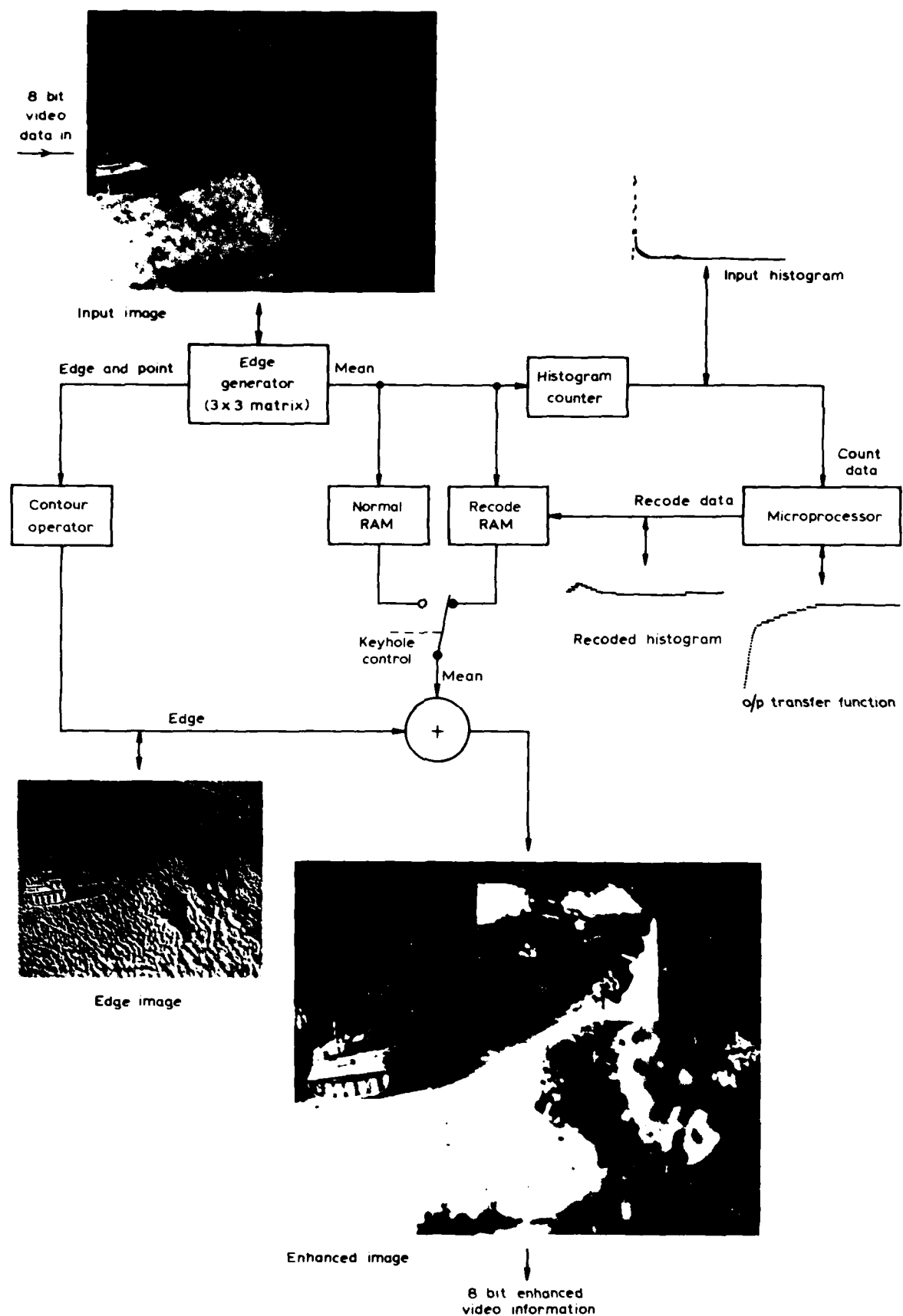
Fig. 6   Simplified diagram of 8 bit enhanced grey-level manipulation process video information

Input histogram

Output histogram

256 grey-levels
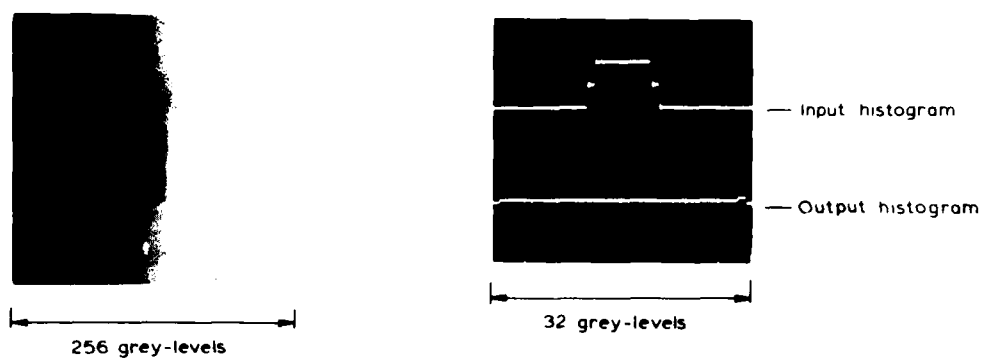
32 grey-levels

Fig. 7 Grey-level histogram equalisation of a 256 step grey-level wedge

Histogram

Histogram

Normal-unenhanced
(a)

Grey-level histogram equalisation
(b)
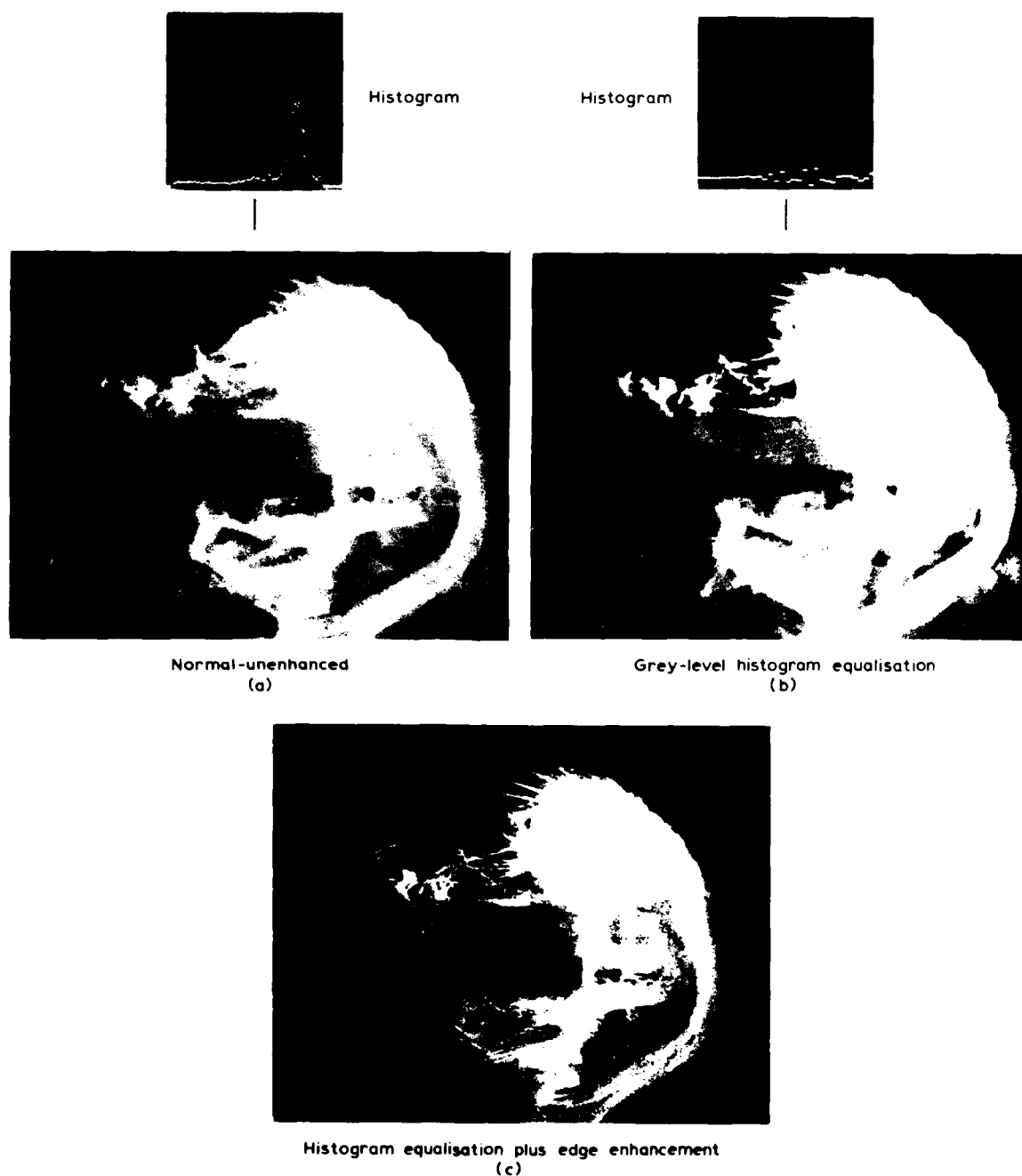

Histogram equalisation plus edge enhancement
(c)

Fig.8 Grey-level histogram equalisation of an x-rayed fish head
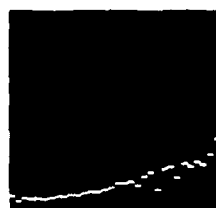
Histogram

Grey-level histogram modification with a rising power

law distribution of $\left(\dfrac{\beta}{4} + j^{\frac{\gamma}{4}}\right)^{*}$

(a)

Histogram

Grey-level histogram modification with a falling power

law distribution of $\left(\dfrac{\beta}{4} + (n-j)^{\frac{\gamma}{4}}\right)^{*}$

(b)

*j is an output level and $o \leq j \leq n$ levels

Fig. 9   Grey-level histogram modification using a power law distribution
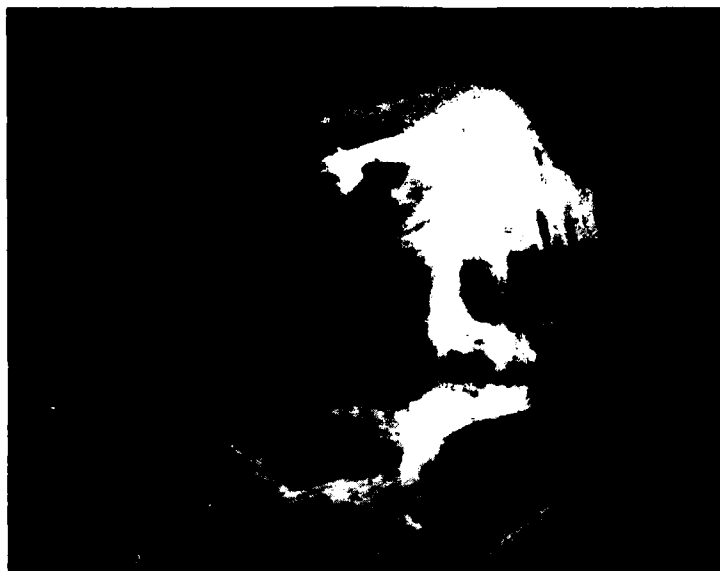
Grey-level histogram equalisation only

(a)

Grey-level histogram equalisation plus 2-dimensional edge enhancement

(b)

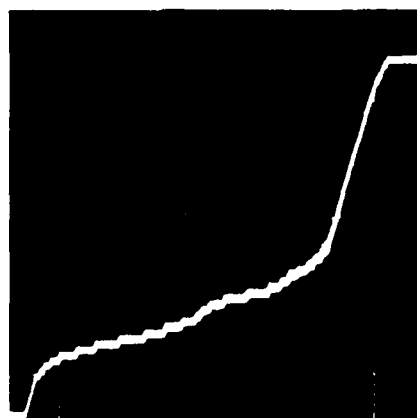Fig. 10   Large area grey-level histogram equalisation

Input histogram

Output histogram

Transfer function

Fig 11 Scene with bi-modal grey-level histogram distribution

TEAMWORK IN TARGET ACQUISITION

by

D. Dey   H. Mutschler
N. Ninz   Fraunhofer-Institut für
Vereinigte Flugtechnische   Informations- und Daten-
Werke GmbH   verarbeitung
Hünefeldstraße 1 - 5   Sebastian-Kneipp-Str. 12 - 14
D-2800 Bremen 1   D-7500 Karlsruhe 1
Germany   Germany

SUMMARY

The performance in target acquisition or real-time reconnaissance can be improved by a team of observers. Different possible team organizations and their characteristics are described. In experiments with forward looking TV-films shot from a low level flying aircraft, different pseudo-team algorithms of one to four operators are considered. The acquisition performances are measured with different criteria and a cost function is evaluated which weights the success and confidence of the acquisition and the time expense.

Contents

1. INTRODUCTION

For air attack RPV's and under special conditions also for reconnaissance RPV's, the typical mission profile could be a fast terrain following flight. For both tasks the RPV would contain forward looking image sensors. Due to the real-time requirements to achieve on-line decisions, the continuous transfer of image data to a ground station via a telemetric data link would be necessary.

The information will be displayed to one or more observers by monitors. Their task includes the detection, classification and identification of targets. This task has to be solved in a short time, for normally the targets have to be acquired (and possibly attacked) during a single approach. The difficulty of acquisition is inherent in the following boundary conditions:

- Narrow field of view resulting from image sensor
- Low image quality due to small data-link bandwidth
- Fast moving scene
- Multiple targets
- Camouflaged targets

As it could been shown in /6/ a single observer cannot cope with these conditions to achieve an acceptable acquisition rate. An approach to increase acquisition performance is to use a well organized team of operators which can overcome the problems of small time budget and high required acquisition rates with a minimum of omissions, false-alarms and wrong classifications or identifications. The definition of an optimized team structure has to be based on a survey of existing literature and task-adequate experiments. The experiments performed here should answer the question, how many team-members in which decision making organization give the best relation success to expense.

2. PRINCIPLES OF TEAM ORGANIZATION

Team organizations can be divided into non-redundant (every operator ob-
serves a different part of the image data) and redundant (all operators
observe the same image data) as it is shown in Figure 1. Non-redundant
teams can share their work in contents or sequence. In case of a con-
tents sharing structure an increase of the operators number increases per-
formance by the same factor /2/. This has not been proved yet for the
division into job-sequences.

For target acquisition by RPV with a single sensor the first approach to
an optimized team structure should be a redundant organization to assure
a minimum of missed and false classified targets. A subdivision of re-
dundant team organization leads to real teams and pseudo-teams determined
by the kind of communication between team-members.

2.1 REAL TEAM WITH REDUNDANT STRUCTURE

A real team of operators cooperates directly to solve the observation task,
a pseudo-team organization connects the results of single observers by an
algorithm without intercommunication.

A real team with non-hierarchical structure has always a better performan-
ce than single observers for errors are eliminated by other team members
/3, 4, 5, 6, 7/. The performance increases with growing team size, the
greatest increase is achieved by adding a second observer to a single one.
Further enlarging of the team leads asymptotically to a certain level in
performance /2, 8/. A team of the described structure is significantly
improving the acquisition especially when task difficulty /9/ and task
frequency /4/ are high.

The real team with hierarchical structure can evolute or can be defined
formally. The genuine team evolutes due to the differing dominance of
team members. A maximum performance of a hierarchical team presumes a be-
haviour of single members between authoritarian and democratic. Total
lack of dominance as well as a significantly dominant leader would decrease
performance. A superior performance can be expected if all team members
are active and their cooperation is organized by a coordinator /10/.

A formally defined hierarchy requires that the most experienced operator
acts as team leader /11/.

2.2 PSEUDO-TEAM WITH REDUNDANT STRUCTURE

This kind of team solves the acquisition task without interpersonal commu-
nication. An algorithm combines the results of the single team members to
a final decision. The algorithm has to be chosen in regard to the desired
performance criteria of the team.

If the observation task leads to a yes-no-decision two sets of algorithm
can be separated. In a serial algorithm the team decision is positive if
all single results are positive. A parallel algorithm produces a positive
decision of the pseudo-team if one single decision is positive (Fig. 2).
The serial algorithm minimizes the false-alarm rate, the parallel algorithm
maximizes the detection rate /12/. The mixture of both algorithms leads
to a partly serial algorithm where a defined number of positive single de-
cisions are necessary to produce a positive decision of the pseudo-team.
The serial as well as the parallel types of algorithms increase detection
rate with increasing number of team members asymptotically /13/.

The final level of detection rate is always lower in the serial case than
in the parallel /8, 14/.

The false-alarm rate increases linear with team-size using a parallel algo-
rithm /13/ whereas the serial one is minimizing this kind of error, so
that independent observers have a better conformity in detecting targets
than confusing objects /12/.

It is also possible to weight the single observers decisions in the pseudo-
team algorithm. This is necessary if a confidence or probability has to
be assigned. No influences on the team performance was found /15/ with
the weightings:

- unique
- not unique, assigned by self-rating
- not unique, resulting from past performance

Real teams and pseudo-teams have some mixed forms with different feedbacks.

- Information about the decision of other team members increases perfor-
  mance of a serial pseudo-team /16/.

- The results of a pseudo-team can be fed back to the single observers and discussed in a real team (team consensus feedback) or they can be displayed one /17/ or several /15/ times to the pseudo-team (consensus feedback). The team consensus feedback reduces the false-alarm-rate /17/.

- Results of a pseudo-team can be observed by the single operators without communication /9/.

A raw comparison of team organizations gives the following ranks in regard to detection rate and speeds /6, 9, 12/:

1) parallel pseudo-team
2) real team
3) single observer
4) serial pseudo-team

The ranks in minimizing false-alarm-rate are

1) serial pseudo-team
2) real team
3) single observer
4) parallel pseudo-team

This survey shows that a complex task as under consideration here has not been investigated in the literature, but basic decisions can be made how to chose a team. The choice has to be made between pseudo-teams and single observers because real teams can be neglected in this real time conditions. In the two rank-scales a pseudo-team gives the best results in both cases but an optimum has to be found between the serial and the parallel type due to the defined aims in target acquisition.

3. EXPERIMENTS

The reported experiments and results are a selection from /18/.

3.1 EXPERIMENTAL SET-UP

Figure 3 shows the experimental set-up. In a simulation cabin of 2,5 m x 5 m four operators are sitting in front of TV-monitors with 23 respective 31 cm diagonals. They watch a TV film replayed by tape recorder which contains fifteen film scenes of 45 - 60 sec duration with a pause picture of 3 sec inbetween. The films were taken during an experimental Remote Target Acquisition Program (KEL/TZE). They were made with a 625 line-camera with zoomable field of view between $4,5^\circ - 45^\circ$ and variable pitch between 3 and 15 degrees.

All operator panels are equipped with light-pens to mark the x/y-coordinates of objects like cars, trucks and tanks and four push buttons for the classification to be operated by the left hand. The operator places were separated by thin walls. All information were fed into the computer for on and off-line evaluation.

3.2 EXPERIMENT ORGANISATION

The experiments were developed during pretests with four operators. In the preliminary trials it became obvious that identification was not possible due to image quality. Therefore the assumption was made that all targets were "enemy" targets implied by tactical situation. The task of the operators was

- to detect and mark coordinates by light pen
- to classify with push-buttons

as quick and correct as possible when they have seen a

- car
- truck or
- tank

The main tests were made with two groups of four operators which were employees of the company with no experience in target acquisition on monitors.

The duration of a test run was 15 minutes. Before a trial a written instruction was given to the operators. Also the handling of the light pen and the push-buttons were trained before each test-run. One to two runs per day were made.

All film scenes contained a total of seventeen targets. Some scenes had more than one, others no target. The performances of each test-run were evaluated and after the ninth run the operators were called trained. The test serie was prolonged until thirteen runs so that four tests for final evaluation are usable.

It was a deficiency of the experiments that only one film was available so that the learning behaviour of the operators is not representative for real situations.

3.3 TV-MATERIAL USED

In /1/ scene parameters as context, target size, contrast etc. are defined. Table 1 shows the regarded scene parameters. For parameters not considered no data were available or could not be used statistically.

Table 2 shows the evaluated parameters of all targets. The criterion value is the mean over eight operators and four trials for each single target.

Figure 5 shows the histograms of target and background brightness and contrast and the mean and $\pm 1\sigma$ values for the three target categories.

Table 3 gives detailed information about scenes parameters and their influence on acquisition measures and criterion function value. Interesting is that there is no significant influence of target motion on acquisition performance as is proved for acquisition with non-moving camera. Here it has to be recognized that the scenes themselves are dynamically and that there were only three fast moving targets.

It can be seen that the target-background contrast has a significant influence on acquisition performance.

Interesting is a comparison between a subjective classification of the difficulty to detect and classify special targets and the defined criterion function J. It can be seen that the J-function value correlates very well with the subjective difficulty, see table 3.

Table 4 gives a qualitative weighting of scene parameters on acquisition performance.

Acquisition performance is improved by target brightness, target background-contrast, civil cars and non-moving targets, it decreases with camouflaged targets like tanks.

3.4 DEFINITION OF MEASURED VALUES

The following definitions were used to describe the target characteristics, the single operator performance and the team performance:

- Display time T_Z:

  Period in which the target is detectable, starting with the earliest time the target could be detected and ends with the disappearance from the screen.

- Contrast K

  $$K = \frac{L_1 - L_2}{L_1}$$

  with L_1: Brightness of the brighter object
  L_2: Brightness of the darker object

- Detection time T_E:

  Period to detect a target correctly and setting coordinates by light-pen, starting with the time the target can be detected.

- Classification time T_K:

  Period between the correct detection and the correct classification of a target (car, truck, tank). The coordinates are taken as correct, if they correspond to the apriori known target position.

- Total time T_G: Period to detect and classify a target correctly

  $$T_G = T_E + T_K$$

The following performance indicator dimensions are %-based on the number of displayed targets (Fig. 4 gives a systematic overview):

Correct Detections $\qquad$ N_{RE}

Missed Detections (Omissions) $\quad$ N_{NE}

False Alarms $\qquad$ N_{FE}

with

$$N_{NE} + N_{RE} = 100\ \%$$
$$0 \leqslant N_{FE} < \infty$$

Correct Classification $\quad$ N_{RK}

Missed Classification $\quad$ N_{NK}

False Classification $\quad$ N_{FK}

with

$$N_{NK} + N_{FK} + N_{RK} = 100\ \%$$

The term decision was used in relation to the pseudo-team performance with the assumption that a correct decision of the team presumes a correct detection and classification of the team members involved.

Correct Decision $\qquad$ N_{RES}

Missed Decision $\qquad$ N_{NES}

False Decision $\qquad$ N_{FES}

with

$$N_{NES} + N_{FES} + N_{RES} = 100\ \%$$

To quantify target acquisition performance of one operator or different teams of operators the performance indices "Success", represented by percent correct classification, "Confidence" by percent false classification and "Speed" by the total time needed in relation to the mean time of targets displayed are combined in a performance criterion J.

The criterion J is defined as

$$J = a\ \bar{N}_{RK} - b\ \bar{N}_{FK} + c(\bar{T}_Z - \bar{T}_G)$$

This empirical criterion function suits for real-time tasks and can be used for the comparison of different acquisition teams. The bars indicate mean values.
It can also be used to look at the influence of different scene-parameters on the acquisition performance.

4. RESULTS

4.1 SINGLE OPERATOR PERFORMANCE

To get a quasi stationary operator behaviour, it was necessary to run the tests thirteen times. Several performance criteria of each operator were calculated. As an example the results for the operator number four are shown in Fig. 6 to 8.

It can be seen that some of the performance measures give redundant information.

Looking at the behaviour of all operators the following observations were made:

- The total time T_G and their variance has a very small tendency to decrease

- The classification time is nearly constant that means it shows more the time to handle the equipment than a decision time. It is thought that the decision is made during the detection process.

- The detection time is responsible for the shape of the total time T_G.

- There is only a small difference between correct detections and correct classifications that means the classification decision was included in the detection decision.

- Wrong classifications are nearly constant over test runs.

- False detections vary very much from test to test. It shows the activation level and the risk behaviour of the operators.

- The tendency can be seen that operators with many false detections also have many right detections.

4.2 CRITERION FUNCTION

Fig. 9 to 11 show the mean values of the last four test-runs of every operator for all performance measures and the mean value of the group.

Fig.12 shows different criterion function values for every operator of group one and two. A variation of the coefficients should give a feeling for the sensitivity of the criterion concerning the weighting of success, confidence and speed.

Which coefficients (a, b, c) have to be taken for each of the three performance factors depend on the aims and criteria of the task.

For the comparison of team algorithm a unique weighting was chosen with (a = 1, b = 1, c = 30).

The rank of the operators in relation to their criterion value (best, second, third performance) depend on the weighting coefficients used.

The result for individually parallel work of all operators is $J = 58,5$ with $\sigma = 14,1$. This is the basis for discussions of improvement of performance by pseudo-teams.

The mean target display time was 2,9 sec with $\sigma = 1,5$. In /1/ it was found that display time of 2 sec resulted in a heavy operator load. This can be confirmed by the tests.

4.3 TEAM DECISION "N OUT OF FOUR"

Four different algorithms are possible. One out of four means: The quickest classification is taken as the decision of the team. The time needed (total time) is the one of this operator (parallel pseuao-team).

Two out of four means: The first two equal classifications are taken as decision of the team (mixed parallel-serial pseudo-team). The time needed is the time of the last of the two operators.

Three out of four is equivalent to two out of four.

Four out of four means: All operators have classified the target identically. The time needed is the one of the slowest operator (serial pseudo-team).

Fig. 13 shows the results for both groups of operators. One out of four resulted in a better performance of about 25 % than the best single operator.

The more operators are involved, the slower is the decision that means the total time goes up (40 %), the number of no decisions goes up (43 %), the number of correct decisions goes down. The positive effect is that the number of wrong decisions goes down (11 %).

Fig. 14 shows the sensitivity of the criterion function for the four pseudo-teams. It can be seen that the algorithm one out of four gives the best performance, if number of targets (success), number of false alarms (confidence) and time needed (speed) are equally weighted.

4.4 TEAM DECISION "ONE OUT OF M "

The evaluation of this algorithm should show the "cost effectiveness" of the number of operators.

A problem was, which operators of a group should be taken for one out of one, one out of two and one out of three. It was decided to take not the best or the worst.

Fig. 15 shows the results for the second team. With the number of operators the time needed (total time) goes down, correct decisions increase, number of no decision decreases, but wrong decisions increase.

The sensitivity of the results as a function of weighting factors for success, confidence and speed is shown in Fig. 16.

It can be stated that one out of four is 25 % and one out of three 14 % better than the best single operator.

The optimum of the criterion function lies between three and four operators depending on their skill and the weighting factors, i. e. the aimes of target acquisition.

5. CONCLUSION

The optimal size of a pseudo-team is three to four.

Assumed that there is an equally weighting of success, confidence and speed the best performance is given, when the first decision of one of the operators is taken (parallel pseudo-team).

The investigated non-hierarchical team structure has a disadvantage, because each operator has to perform two tasks, detecting targets by searching the possible area of appearance and tracking of targets until a classification is possible. If there are targets in a short sequence, it can happen that they are over-looked. This can be improved by a hierarchical team.

It is assumed that with image enhancement techniques such as gray scale manipulations, the acquisition performance can be improved.

In the future it should also be investigated how a priori information (like high altitude reconnaissance) will improve the operator performance.

Additionally the influence of an automatic detection device on the total acquisition performance should be investigated.

REFERENCES

/1/ Mutschler, H.:
Fraunhofer Institut für Informations- und Datenverarbeitung, Karlsruhe;
LEISTUNGSFÄHIGKEIT EINES FERNSEHBEOBACHTERS BEI ECHTZEITAUFKLÄRUNG;
Forschungsvorhaben T/RF 36/71530/71334;

/2/ Kidd, J.S.:
A COMPARISON OF ONE-, TWO-, AND THREE-MAN WORK UNITS UNDER VARIOUS CONDITIONS OF WORK LOAD;
Journal of Applied Psychology, Vol. 45, No. 3, 1961, 195 - 200;

/3/ White, C.T.;
Ford, A.:
EYE MOVEMENTS DURING SIMULATED RADAR SEARCH;
Journal of the Optical Society of America, Vol. 50, No. 9, Sept. 1960;

/4/ Wiener, E.L.:
INDIVIDUAL AND GROUP DIFFERENCES IN INSPECTION;
in: HUMAN RELIABILITY IN QUALITY CONTROL;
Drury, C.G.;Fox, J.G. (Ed.);
London; Taylor & Francis Ltd.; 1975;

/5/ Zajonc, R.B.:
SOCIAL FACILITATION;
Sience, 1965, 149, 269 - 274;

/6/ Morrissette, J.O.;
Hornseth, J.P.;
Shellar, K.:
TEAM ORGANIZATION AND MONITORING PERFORMANCE
Human Factors, Vol. 17, No. 3, 1975, 296 - 300;

/7/ Johnston, W.A.;
Briggs, G.E.:
TEAM PERFORMANCE AS A FUNCTION OF TEAM ARRANGEMENT AND WORK LOAD;
Journal of Applied Psychology, Vol. 52, No. 2, 1968, 89 - 94;

/8/ Doten, G.W.;
Cockrell, J.T.:
U.S. Army Personnel Research Office;
THE USE OF TEAMS IN IMAGE INTERPRETATION: INFORMATION EXCHANGE, CONFIDENCE, AND RESOLVING DISAGREEMENTS;
Oct. 1966, Technical Research Report 1151;

/9/ Bolin, S.F.;
Sadacca, R.;
Martinek, H.:
U.S. Army Personnel Research Office;
TEAM PROCEDURES IN IMAGE INTERPRETATION;
Dec. 1965, Technical Research Note, 164;

/10/ Sandner, D.:
PSYCHODYNAMIK IN KLEINGRUPPEN;
1978;

/11/ Siebel, A.:
Carl-Cranz-Gesellschaft e. V., Heidelberg;
PROBLEME DER MILITÄRISCHEN LUFTBILDAUSWERTUNG;
März 1973, Vortrag im Lehrgang OP3 "Luftaufklärung";

/12/ Sadacca, R.;
Martinek, H.;
Schwartz, A.I.:
U.S. Army Personnel Research Office;
IMAGE INTERPRETATION TASK-STATUS;
Oct. 1962, Technical Research Report 1129;

/13/ Williams, L.G.:
TARGET CONSPICUITY AND VISUAL SEARCH;
Human Factors, 8, 1960, 80 - 92;

/14/ Green, D.M.;
Swets, J.A.:
SIGNAL DETECTION THEORY AND PSYCHOPHYSICS;
Ed.: Wiley and Sons, Inc.;
New York, London, Sydney, 1966;

/15/ Staël von Holstein, C-A.S.:
ASSESSMENT AND EVALUATION OF SUBJECTIVE PROBABILITY DISTRIBUTIONS;
The Economical School of Economics;
Stockholm, 1970;

/16/ Beechler, R.L.;
Winterstein, S.H.;
Kamper, R.M.;
Jeffrey, T.E.:
U.S. Army Behavioral Science Research Laboratory;
A STUDY OF RAPID PHOTOINTERPRETATION METHODS;
June 1969, Technical Research Report 1153;

/17/ Cockrell, J.T.:
U.S. Army Behavioral Science Research Laboratory;
MAINTAINING TARGET DETECTION PROFICIENCY THROUGH TEAM CONSENSUS FEEDBACK;
Dec. 1969, Technical Research Note 219;

/18/ Dey, D.;
Finke, E.;
Ninz, N.;
Pittwohn, H.:
Vereinigte Flugtechnische Werke, Et41, Bremen;
UNTERSUCHUNG DER ZIELAKQUISITIONSLEISTUNG EINES TEAMS BEI AUTOMATISCHER
ZIELERKENNUNG;
March 1980, Endbericht zum Forschungsauftrag T/RF36/71528/71332.

List of possible scene parameters	Parameters considered
Target size	resulting from target shape
Number of targets per TV-scene	1-6 (identical targets)
Shape of target	car, truck, tank
Target brightness	yes
Target color	-
Number of target categories	3
Target grouping	-
Target motion	yes
Background brightness	yes
Target-background contrast	yes
Context of background	yes
Target density	-
Target position, traces	-
Environmental parameters	-

Table 1: Scene Parameters

Target No.	Class	Context	Movement	Difficulty	L_z ($\frac{cd}{m^2}$)	L_H ($\frac{cd}{m^2}$)	K (s)	T_Z (s)	T_g (s)	Z_{RK} (%)	Z_{FK} (%)	J (a=1, b=1, c=30)
1	CAR	Forest	fast	medium	6	0,7	89	2,3	2,3	3,1	0	38
2	4 TANK	Heather	no	difficult	6,8	8,6	35	4,5	3,8	78	47	-25
3	CAR	Trees	no	easy	9,4	4,5	52	3,0	2,4	100	3	94
4	TRUCK	Street	slow	easy	10,2	2,7	73	2,2	1,6	100	0	135
5	CAR	Trees	no	difficult	11,8	3,0	75	0,8	1,7	47	0	46
6	CAR	Meadow	no	medium	15,6	8,3	47	2,9	2,8	34	0	19
7	TRUCK	Trees	no	medium	6,4	4,4	31	1,5	2,0	100	0	113
8	TANK	Trees	fast	easy	18,8	7,2	62	2,8	2,2	100	3	121
9	5 TANK	Heather	no	difficult	1,6	1,9	18	7,3	5,2	88	22	-43
10	TRUCK	Trees	slow	difficult	4,4	8,1	46	2,0	2,7	6	3	3
11	TRUCK	Trees	slow	medium	8,8	2,2	75	4,6	5,3	97	0	13
12	TRUCK	Trees	fast	difficult	16,7	11,1	34	4,0	2,6	34	59	-23
13	6 TANK	Heather	no	easy	5,2	11,0	53	3,1	1,2	97	3	111
14	TANK	Heather	no	easy	13,3	1,7	79	2,7	2,1	88	6	81
15	TRUCK	Heather	no	medium	6,7	0,7	63	2,5	2,0	59	6	72
16	TRUCK	Heather	slow	difficult	1,6	5,1	29	1,5	0,9	9	9	49
17	TRUCK	Heather	no	difficult	0,9	3,2	93	1,6	2,1	22	3	50

Table 2: Characteristic data and performance
measures of all targets

Scene Parameters/Acquisition Performance	TARGET CATEGORIE			TARGET MOTION			CONTEXT OF BACKGROUND			RATED DIFFICULTY TO CLASSIFY TARGET		
	car	truck	tank	no	slow	fast	heather	trees	other	easy	medium	difficult
Number of targets	5,00	8,00	4,00	10,00	4,00	3,00	7,00	7,00	3,00	5,00	5,00	7,00
Mean target brightness $L_z \left[\frac{cd}{m^2}\right]$	12,30	6,90	6,70	7,80	6,30	13,80	5,20	10,90	10,60	11,40	8,70	6,30
background brightness $L_H \left[\frac{cd}{m^2}\right]$	4,70	4,70	5,80	4,73	4,50	6,30	4,60	8,00	5,50	5,40	3,20	5,90
target-background contrast $[K]$	65,00	55,00	46,30	54,60	55,80	61,60	52,90	53,60	69,60	63,80	6,10	47,10
time of target detectable T_z $[s]$	2,36	2,40	4,40	3,00	2,60	3,00	3,30	2,67	2,46	2,80	2,80	3,10
total time $[s]$	2,28	2,40	3,00	2,50	2,60	2,40	2,47	2,70	2,20	1,90	2,90	2,70
correct classification z_{RK} $[\%]$	62,40	53,30	87,80	71,30	53,00	55,00	63,00	69,00	55,00	97,00	64,20	40,50
false classification z_{FK} $[\%]$	1,30	9,40	19,50	9,00	1,50	20,60	12,90	9,70	-	3,00	1,20	19,60
criterion J (a = 1, b = 1, c = 30)	63,60	51,50	31,00	51,80	50,00	45,30	42,10	52,40	64,00	108,40	51,00	8,14

Table 3: Scene parameters and acquisition performance
in regard to target characteristics

Scene parameters	Effect on target acquisition
High target brightness	+
High background brightness	O
High Target-background contrast	+
Context of background	O
Target categorie: car	+
truck	O
tank	-
Target motion: no	+
slow	O
fast	-

+   increase of acquisition performance

O   neutral

-       decrease of acquisition performance

Table 4: Qualitative weighting of scene parameters
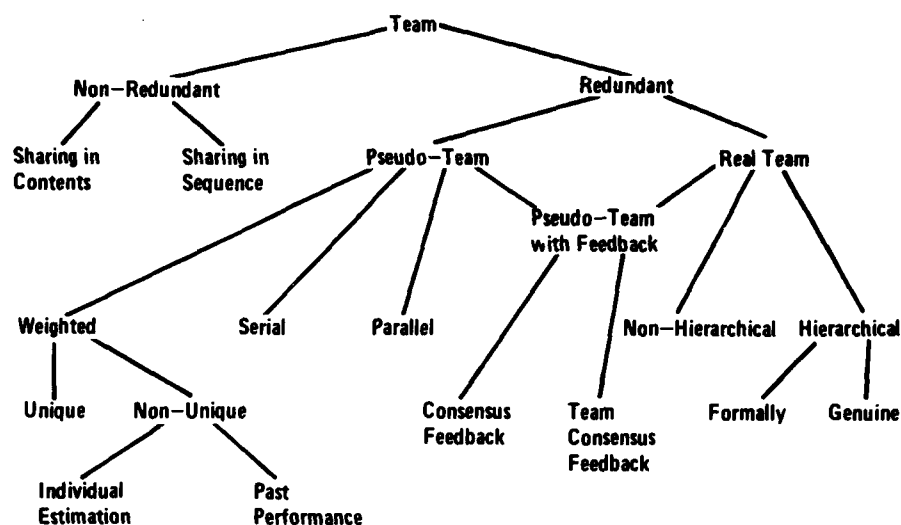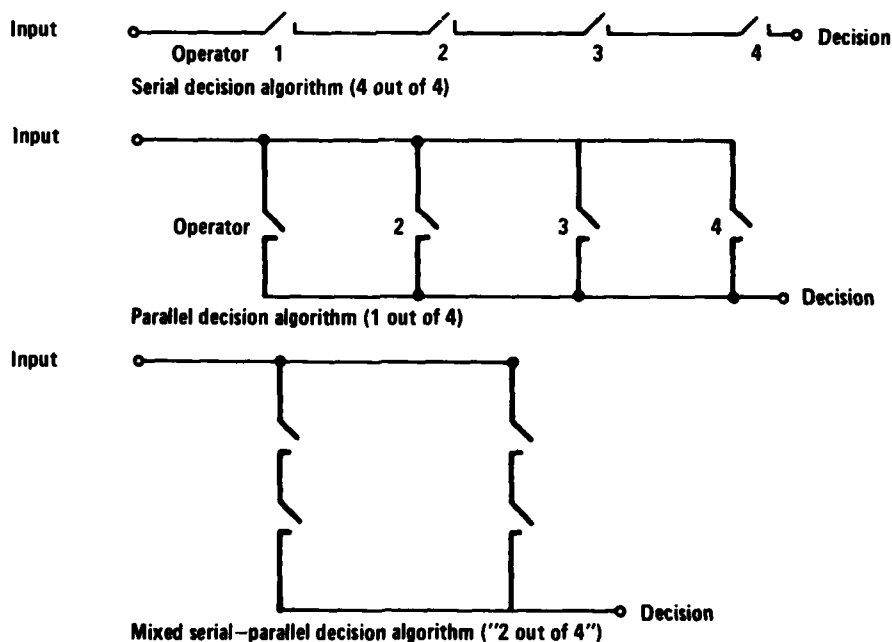for target acquisition

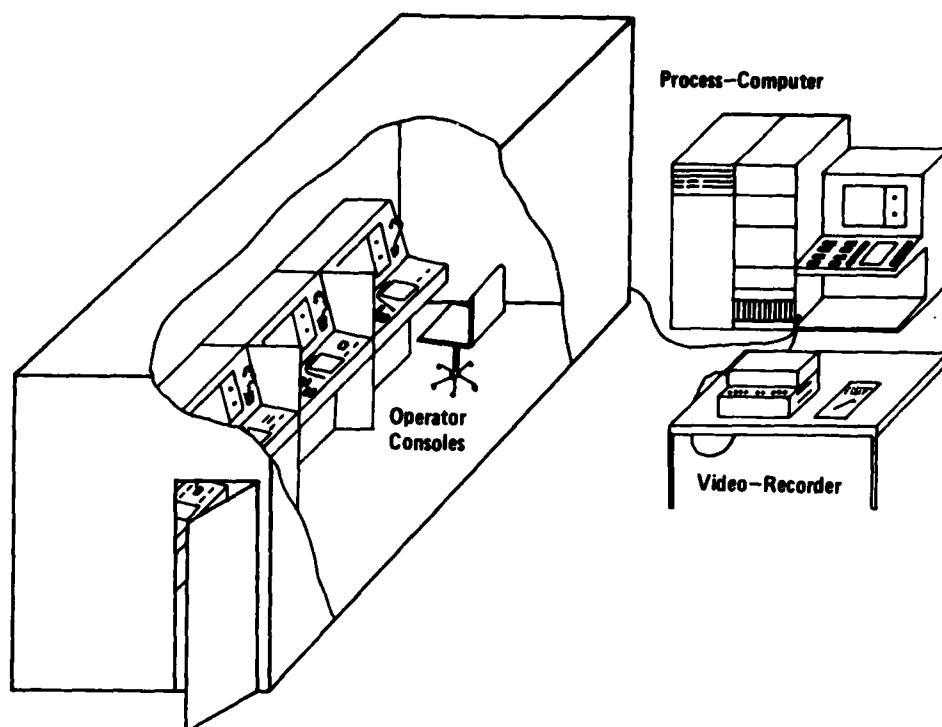Fig. 1   Types of team organisations

Serial decision algorithm (4 out of 4)

Parallel decision algorithm (1 out of 4)

Mixed serial–parallel decision algorithm ("2 out of 4")

Fig. 2   Pseudo-team algorithms

Fig. 3 Experimental set-up

TIME MEASURES

INFORMATION
COLLECTED

* Time target is
  displayed T_Z

Target leaves screen

* Total time T_G
  sum of T_E and T_K

* Classification time T_K
  for correctly classified
  targets (difference be-
  tween T_E and actual
  time)

Operators classifying target
car, trucks, tanks

* Target class (car, truck
  tank)
  – correct classifications
  – false classifications
  – non-classifications

* Detection time T_E for
  correctly detected targets
  (difference between T_O
  and actual time)

Operators searching
for target

* Object coordinates
  – correct detections
  – false alarms
  – missed targets

T_O

Target detectable

* Target coordinates

Fig. 4 Sequence and data collection in target acquisition
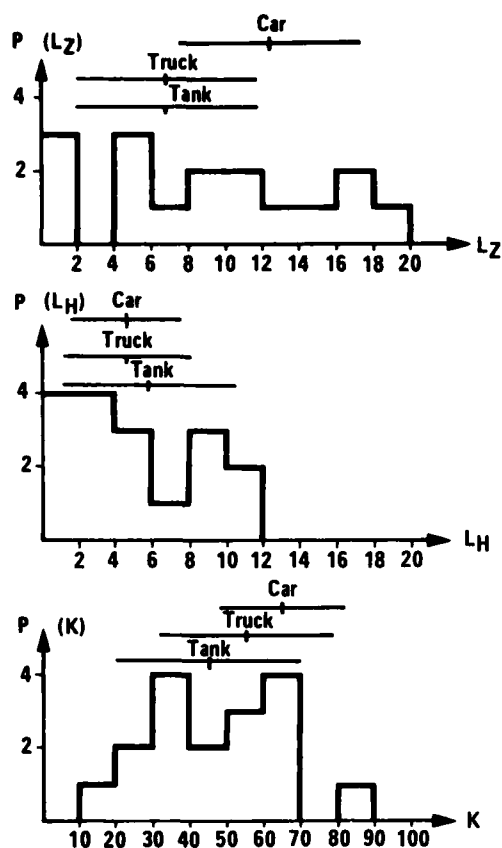
Fig. 5 Histograms of target brightness L_Z, background brightness L_H and contrast K
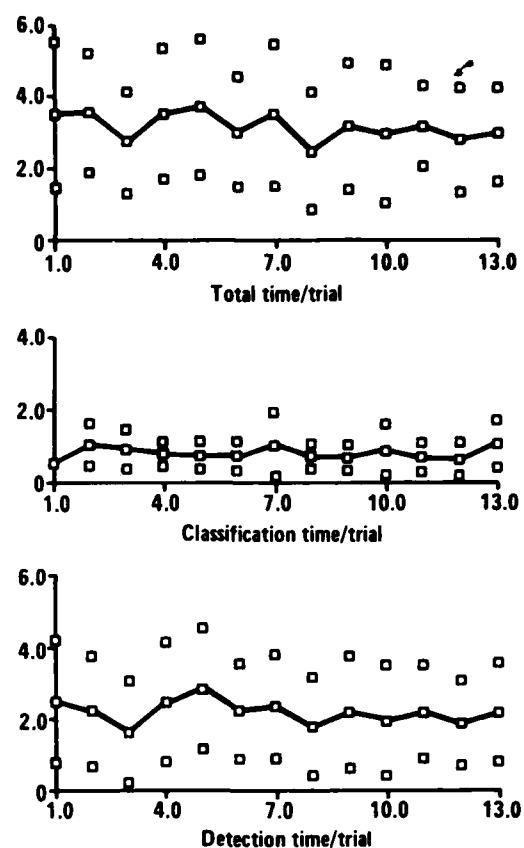
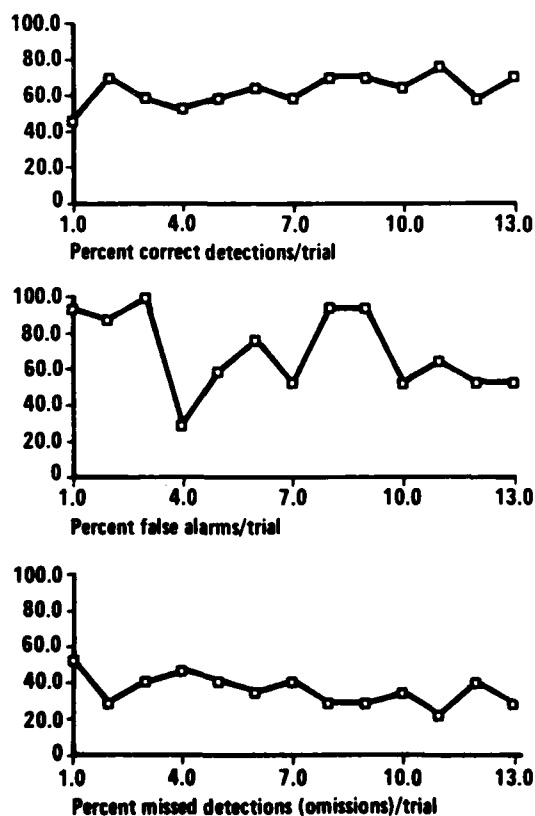Fig. 6 Time performance measures of test person 4

Fig. 7 Detection performance measures of test person 4
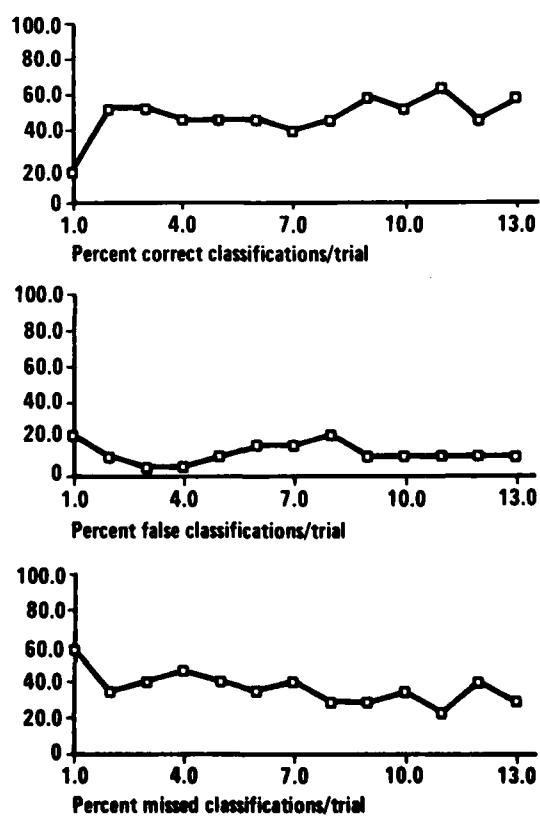

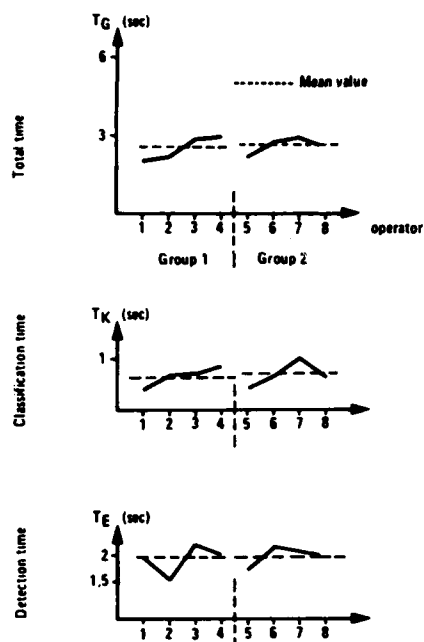Fig. 8 Classification performance measures of test person 4

Fig. 9 Mean time measures of all operators averaged
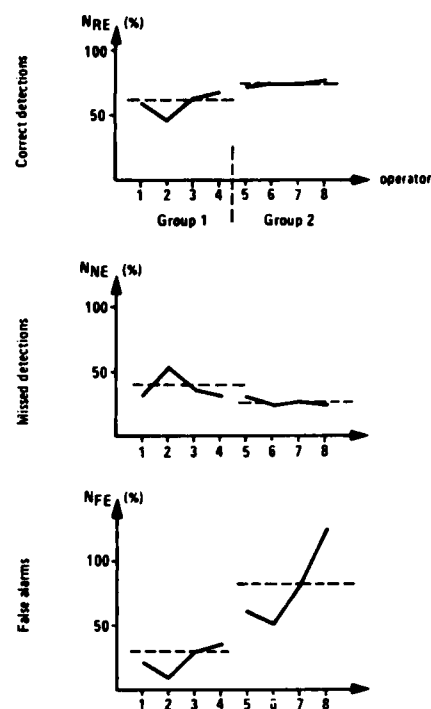over the trials 10 – 13

Fig. 10 Mean detection performance of all operators
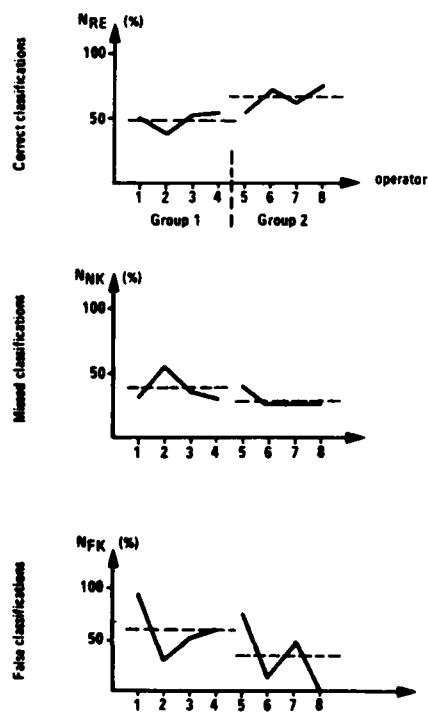averaged over the trials 10 – 13

Fig. 11 Mean classification performance of all operators
averaged over the trials 10 – 13

J_1:a=3,b=1,c=10
J_2:a=2,b=1,c=10
J_3:a=1,b=1,c=10
J_4:a=1,b=1,c=30
J_5:a=1,b=1,c=20
$J_3 = J_6$:a=1,b=1,c=10
$J_3 = J_7$:a=1,b=1,c=10
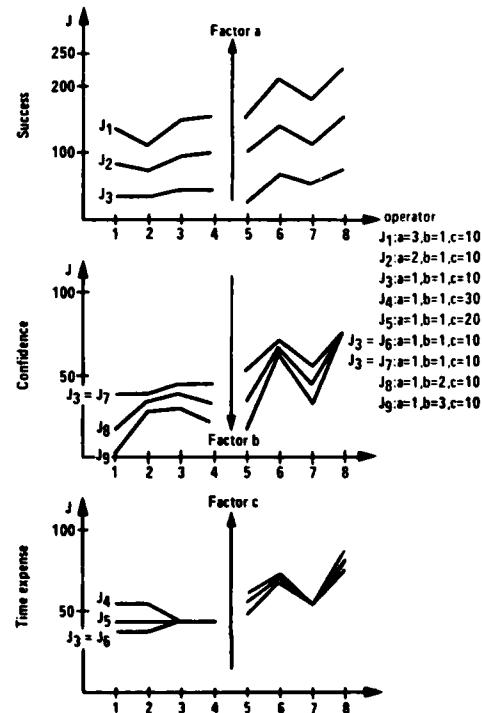J_8:a=1,b=2,c=10
J_9:a=1,b=3,c=10

Fig. 12 Sensitivity of cost function to the value of
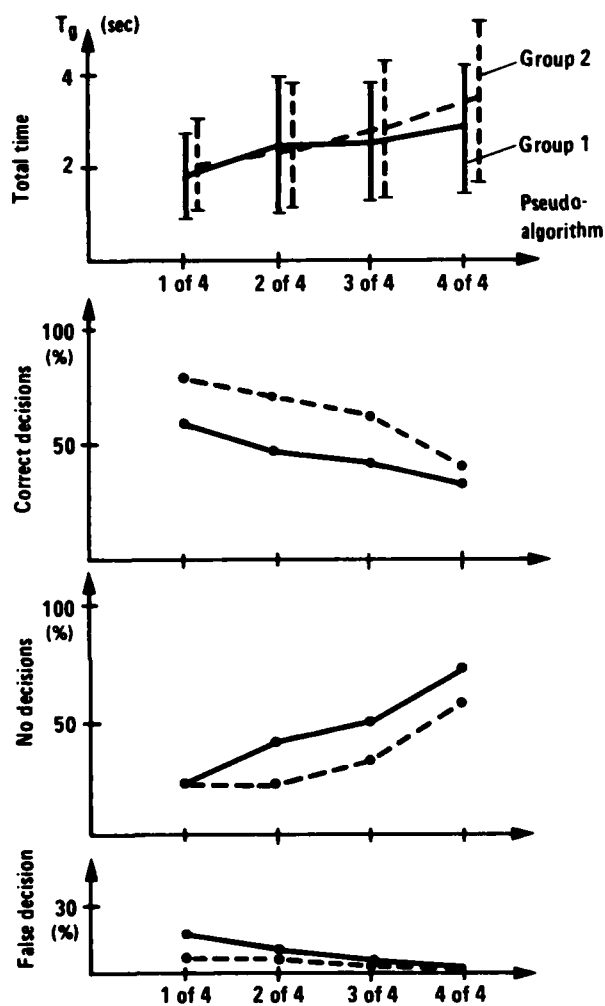weighting coefficients for single operators

Fig. 13 Mean performance of pseudo-teams "N out of 4" averaged over the trials 10 – 13
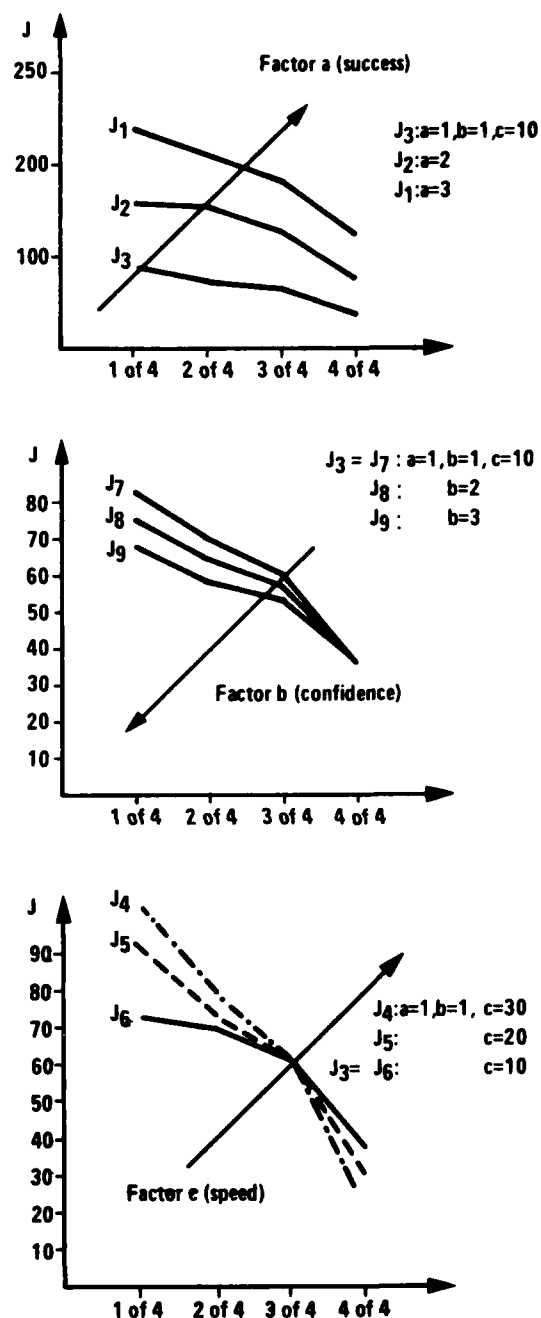

Fig. 14 Sensitivity of cost function to the value of weighting coefficients for "N out of 4" with group 2
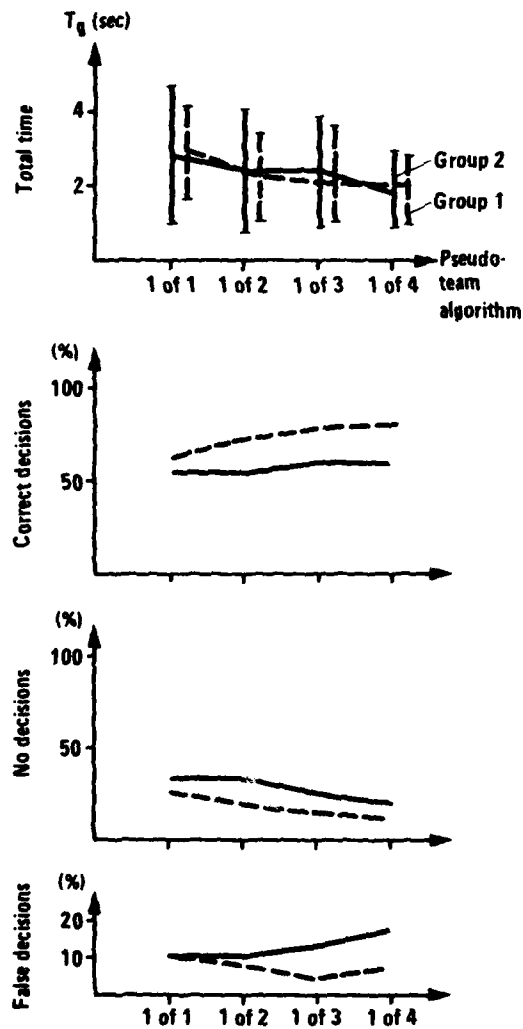
T_g (sec)

Total time

Group 2

Group 1

1 of 1   1 of 2   1 of 3   1 of 4   Pseudo-team algorithm

(%)

Correct decisions

100

50

(%)

No decisions

100

50

(%)

False decisions

20

10

1 of 1   1 of 2   1 of 3   1 of 4

Fig. 15   Mean performance of pseudo-teams "1 out of M" averaged over the trials 10 — 13

J

250

Factor a (success)

$J_1: a=3, b=1, c=10$

200

$J_2: a=2$

100

$J_3: a=1$

1 of 1   1 of 2   1 of 3   1 of 4

J

$J_7 = J_3 : a=1, b=1, c=10$

80

$J_8 : b=2$

70

$J_9 : b=3$

60

50

40

Factor b (confidence)

30

20

10

1 of 1   1 of 2   1 of 3   1 of 4

J

100

$J_4 : a=1, b=1, c=30$

$J_5 : c=20$

90

$J_6 = J_3 : c=10$

80

70

60

Factor c (speed)
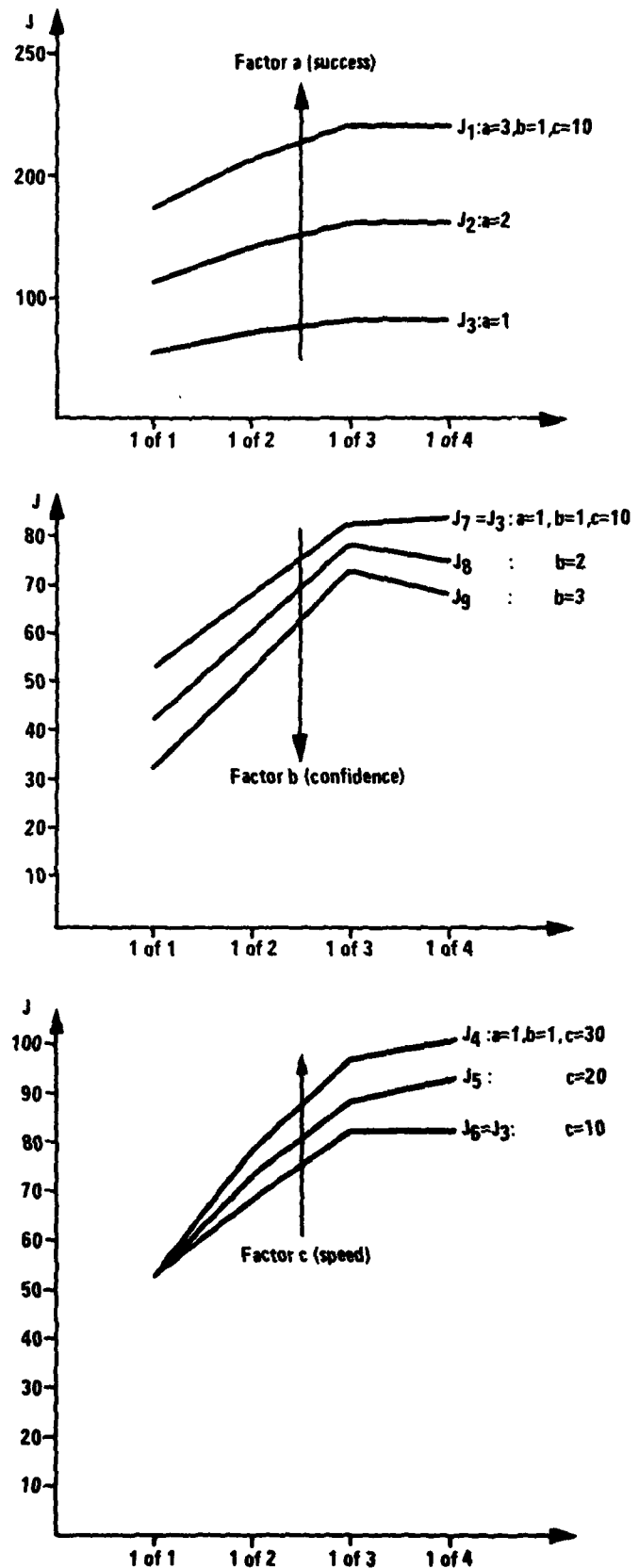
50

40

30

20

10

1 of 1   1 of 2   1 of 3   1 of 4

Fig. 16   Sensitivity of cost function to the value of weighting coefficients for "1 out of M" with group 2

IMAGE PROCESSING TECHNIQUES USING SPLINE APPROXIMATION

by
W.C. Huisman and J. van Kasteel

National Aerospace Laboratory NLR
Anthony Fokkerweg 2
Amsterdam
The Netherlands

ABSTRACT

A two-dimensional data-compression method is described which is based on a least-square image approxima-
tion with use of splines. A complete analysis of this process in the frequency domain is given. For the data-
compression process the feasibility of some candidate realisations of the real-time operating hardware is
described. The image reconstruction from the compressed data set consists of an off-line computation of a modi-
fied compressed data set followed by a real-time replay interpolation process.
In order to achieve a better image reconstruction a flexible edge enhancement algorithm was developed.

1. INTRODUCTION

Limiting the amount of sensor data to be transmitted and/or processed may considerably impact on the total
end-to-end sensor system design. Especially with imaging type of sensors, data reduction techniques may be
beneficially applied, due to the redundancy in the images.
At the National Aerospace Laboratory, data reduction techniques based on spline approximation are inves-
tigated for application to:
- remote sensing systems (LANDSAT)
- sensor systems (IR, SLAR)
- airborne data acquisition systems.
Spline approximation techniques achieve data reduction by entropy reduction in the spatial domain. Atten-
tion to these techniques was drawn because of the attractive local character of the operations to be performed,
and further because of the optimal interpolation property of splines.
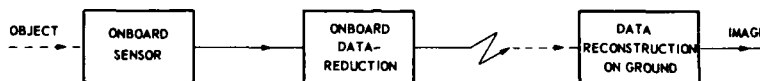

Fig. 1.1 Application of data reduction and reconstruction in an airborne or spaceborne sensor system

The results so far are so promising that application is planned in the data processing of the Dutch Infra-
red Astronomical Satellite IRAS.
These techniques are expected to be very useful as well for reduction of data to be processed for target
acquisition and recognition. In this case one may think of reducing the data from a sensor system onboard the
aircraft, in order to reduce the bandwidth needed for transmission or the capacity of a storage medium. In the
sequel a data reduction method is described which is based on least squares approximation of the image with a
two-dimensional spline function. This spline function is completely defined by a limited number of characteris-
tic parameters. This number is less than the original number of pixels, so data reduction can be achieved by
calculating the characteristic parameters of the approximating spline out of the original pixel values. The
image can be reconstructed from this compressed data set.
The data-reduction process associated with the application of the spline approximation is a simple convolu-
tion filter process, yielding the compressed data set. Image reconstruction is achieved by first preprocessing
the compressed data, which yields a modified compressed data set, and thereafter interpolating this data on a
(variable) grid. The preprocessing is a complex deconvolution process. The interpolation process, however, is
simple again, because only a finite number of values from the modified compressed data set (depending on the
order of the spline) are needed for the computation of the intermediate points. A complete analysis of this
data reduction-reconstruction process in the frequency domain is given in chapter 3 (Renes, J.,79).
Data reduction with this method results in a moderate decrease of spatial resolution. In case of an air-
borne scanner system this is ofter acceptable because of strong correlations between adjacent pixels (over-
sampling in line direction and overlap of scanline) resulting in a resolution sufficient for target acquisi-
tion and recognition. If the decreased spatial resolution is not acceptable some improvement can be obtained
by means of an edge enhancement technique. Further, resolution differences in flight and cross track direction
can be accounted for, by using different reduction factors in both directions.
In view of the airborne application and the data rates involved a hardware implementation of the compres-
sion scheme is of interest. In chapter 4 a detailed treatise is given of some candidate hardware implementa-
tions, for the two-dimensional data-reduction process. Also some remarks on the applicability of the proposed
data reduction-reconstruction process are given. In chapter 5 attention is given to the theoretical backgrounds
of an edge enhancement algorithm. Finally, results obtained with Landsat data and some concluding remarks are
given in chapter 6 and 7.

2. BASIC CONCEPTS

2.1 Splines and B-splines defined

A spline is defined as a function, $S(t)$ say, that results from piecing together polynomial arcs of order
k (degree k-1) as follows

- each arc is defined over a unit length interval; the interval end points are called the knots,
- at each knot the arcs match in value (except for k=1), also the derivatives up to k-1 match then in value.
- for even (odd) k the knots are at (half-) integer values of t.

The B-spline (B stands for Basic) $B_k(t)$ of order k is defined recursively by convolutions as

$$B_1(t) = 1 \text{ when } |t| < \tfrac{1}{2} \text{ and } 0 \text{ elsewhere} \qquad (2.1)$$

$$B_k(t) = (B_{k-1} * B_1)(t)$$

$$= \int^D B_{k-1}(s) \, B_1(t-s)ds \qquad (2.1)$$

$$= \int_{-\frac{1}{2}}^{\frac{1}{2}} B_{k-1}(t-s)ds$$

Consequently $B_k(t)$ satisfies the following properties

(a) $B_k(t) \equiv 0 \quad |t| > k/2$

(b) $B_k(t)$ is a spline of order k $\qquad\qquad (2.2)$

(c) $B_k(t)$ is positive and has unit area

(d) $\sum_\nu B_k(t-\nu) = 1$, all t

Explicit expressions for the first three B-splines are

$$B_1(t) = 1 \quad , \quad |t| < \tfrac{1}{2}$$
$$B_2(t) = 1-|t| \quad , \quad |t| < 1$$
$$B_3(t) = \begin{cases} \tfrac{3}{4} - t^2, & |t| < \tfrac{1}{2} \\ \tfrac{1}{2}(\tfrac{3}{2} - |t|)^2, & \tfrac{1}{2} < |t| < \tfrac{3}{2} \end{cases}$$

The name B(asic)-spline stems from the following representation property:
Any spline S(t) can be represented by the B-spline series

$$S(t) = \sum_\nu C_\nu \, B_k(t-\nu) \qquad (2.3)$$

If S(t) is defined over n intervals the number of coefficients C_ν is n+k-1.
Note that due to (2.2-a) only k consecutive coefficients enter in the evaluation of the series (2.3); e.g. for k=3 one has for $|\epsilon| < \tfrac{1}{2}$

$$S(\nu+\epsilon) = C_{\nu-1}\tfrac{1}{2}(\tfrac{1}{2}+\epsilon)^2 + C_\nu(\tfrac{3}{4}-\epsilon^2) + C_{\nu+1}\tfrac{1}{2}(\tfrac{1}{2}-\epsilon)^2$$

Hence the B-spline coefficients are a sequential, and local representation of the spline S(t).

The notion of B-splines can be generalised to two-dimensional functions using Cartesian products. Instead of knots one now has a grid of lines t=ν and s=μ. The B-splines series then becomes

$$S(t,s) = \sum_\nu \sum_\mu D_{\nu\mu} \, B_k(t-\nu) \, B_k(s-\mu) \qquad (2.3a)$$

In this case only k^2 coefficients enter in the evaluation of the two-dimensional spline function. The function $B_3(t) \, B_3(s)$ is displayed in figure A1.

2.2 Least squares approximation

Data reduction can be achieved by approximating a given function X(t) with a spline S(t). In general the function X(t) is described by a set of samples while the spline S(t) is described by fewer spline coefficients. There are many ways to specify a computation rule for the B-spline coefficients to make the resulting spline S(t) "resemble" a given input function X(t). In least-squares approximation the coefficients are determined by requiring the total power in the error curve X(t)-S(t) to be minimal.

$$\underset{C}{\text{MIN}} \; E_{tot}(C_\nu) = \int (X(t) - \sum_\nu C_\nu \, B_k(t-\nu))^2 \, dt \qquad (2.4)$$

For convenience only the continues case is treated, the discrete case leads to analogue formulations.
By scaling the time t it is no restriction having the spline knots equally spaced with distance 1.
A system with linear equations ensues from setting to zero all partial derivatives

$$\partial E_{tot}/\partial C_\nu = 0 \qquad (2.5)$$

The system, also called the normal equations, has the typical equation

$$\int X(t) \, B_k(t-\nu)dt = \sum_\mu \int B_k(t-\nu) \, B_k(t-\mu)dt \; C_\mu$$

$$= \sum_\mu (B_k * B_k)(\nu-\mu) \; C_\mu \qquad (2.6)$$

$$= \sum_\rho B_{2k}(\rho) \; C_{\nu+\rho}$$

In the last step the symmetry and auto-convolution properties of the B-splines are used.
Before solving the normal equations, the left part, also called the convolutions, has to be computed from the input data X(t). In appendix A, two methods are given for solving the normal equations. When the normal equations (2.6) are solved, the spline S(t) can be computed from equation (2.3).
The analogue to (2.6) for image approximation, with image function X(t,s) is

$$\iint X(t,s) \, B_k(t-\nu) \, B_k(s-\mu) \, dtds = \sum_\rho \sum_\tau B_{2k}(\rho) \, B_{2k}(\tau) \, D_{\nu+\rho, \mu+\tau} \qquad (2.7)$$

Before solving equation (2.7) the left-hand part of this equation, denoted by $Q_{\nu,\mu}$ has to be computed from the input data $X(t,s)$. This computation can be done in two consecutive steps, first compute the convolution $P_\nu(s)$ in t-direction

$$P_\nu(s) = \int X(t,s) \; B_k(t-\nu)dt$$

next compute the final convolution $Q_{\nu,\mu}$ from $P_\nu(s)$

$$Q_{\nu,\mu} = \int P_\nu(s) \; B_k(s-\mu)ds$$

By inspection of (2.6) and (2.7) it follows that (2.7) can be solved in two consecutive steps; first compute $\tilde{D}_{*,\mu}$ from

$$\iint B_k(t-\nu) \; B_k(s-\mu) \; X(t,s) \; dtds = \sum_\rho B_{2k}(\rho) \; \tilde{D}_{\nu+\rho,\mu}$$

and next compute $D_{\nu,*}$ from

$$\tilde{D}_{\nu,\mu} = \sum_\tau B_{2K}(\tau) \; D_{\nu,\mu+\tau}$$

This property is known as the decomposition thereom. When the normal equation (2.7) are solved, the spline $S(t,s)$ can be computed from equation (2.3a).

3. FREQUENCY ANALYSIS

In this chapter an expression is found for the Fourier transform of the spline $S(t)$ in terms of the Fourier transform of $X(t)$ for the least squares approximation.
Analogue expressions can be found for the two-dimensional case. For the frequency analysis we consider $X(t)$ to be defined on $\infty < t < \infty$.
The Fourier Transform of a function $X(t)$ is denoted by

$$FT(X) = x(f) \stackrel{D}{=} \int X(t) \; e^{-2\pi i f t}dt$$

and conversely

$$X(t) = \int x(f) \; e^{+2\pi i f t}df$$

Since

$$FT(B_1) = \sin(\pi f)/\pi f \stackrel{D}{=} \text{sinc}(f),$$

one has from (2.1)

$$FT(B_k) = \text{sinc}^k(f).$$

Sampling of a function $X(t)$ (which will be denoted by $X^*(t)$) at integer values amounts to multiplication with the Dirac comb $III(t)$ defined by the series

$$III(t) = \sum_\nu \delta(t-\nu),$$

where $\delta(t)$ is the Dirac delta function. $III(t)$ has the curious property that it is its own Fourier Transform

$$FT(III) = III(f) \tag{3.2}$$

The convolution of two functions $X(t)$ and $Y(t)$ is denoted by the asterisk "*"

$$(X*Y)(t) = \int X(s)Y(t-s)ds,$$

and by the convolution theorem one has

$$FT(X*Y) = x(f)y(f) \tag{3.3}$$

Using (3.2) and (3.3) one has

$$FT(X.III) = x(f) * III(f) = \sum_n x \; (f-n) \tag{3.4}$$

which co-incides with the periodic repetition of the spectrum $x(f)$.
If $X(t)$ is band-limited to the Nyquist range $-\tfrac{1}{2}<f<\tfrac{1}{2}$, the replications do not overlap.
From the previous discussion it follows that the process in determining the approximating spline $S(t)$ in the least squares sense, can be described with the following system configuration.
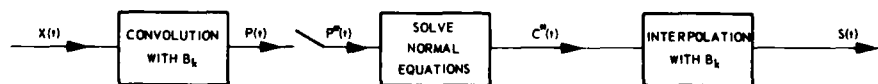

Fig. 3.1 Determination of the least square approximation

The symbol between $P(t)$ and $P^*(t)$ stands for the sampling mechanism.
In the sequel the transfer function of each system component is analyzed separately and, by applying the convolution theorem, the relation between $s(f)$ and $x(f)$ is obtained.

a. Convolution and sampling

The relation between $X(t)$ and $P(t)$ is given by

$$P(t) = (X*B_k)(t)$$

With help of the convolution theorem and (3.1) it follows

$$p(f) = x(f). \; b_k(f)$$
$$= x(f). \; \text{sinc}^k(f)$$

For $k=3$, the function $\text{sinc}^3(f)$ is plotted in figure A.2.
Next, applying the sampling theorem gives

$$p^*(f) = \sum_n x(f-n) \, \text{sinc}^k(f-n)$$

<u>b</u>. Solving the normal equations

The relation between the sampled convolution $P^*(t)$ and sampled B-spline coefficients $C^*(t)$ is given by

$$P^*(t) = (B^*_{2k} * C^*)(t)$$

Applying the convolution theorem from equation (2.4) gives

$$p^*(f) = b^*_{2k}(f) \, c^*(f)$$

Thus

$$c^*(f) = \frac{1}{\sum_n b_{2k}(f-n)} \, p^*(f) \overset{\text{n}}{=} \frac{1}{d_k(f)} \; .$$

For $k=1,2$ and 3 the inverse transfer function $d_k(f)$ of the deconvolution process is given explicitly below.

$$d_1(f) = 1$$
$$d_2(f) = 1 - 2/3 \sin^2 \pi f$$
$$d_3(f) = 1 - \sin^2 \pi f + 2/15 \sin^4 \pi f$$

In general the periodic function $d_k(f)$ can be computed from (using eq. (3.7)).

$$d_k(f) = \sum_n B_{2k}(n) \, e^{-i2\pi n f}$$

The actual relation between $P^*(t)$ and $C^*(t)$ is determined by the chosen implementation method, see Appendix A. In case the Cholesky decomposition method is used, the relation between $p^*(f)$ and $c^*(f)$ almost equals the ideal relation (3.8).
The periodic transfer function $d_3^{-1}(f)$ is plotted in figure A3. When the local deconvolution process is used conform equation (A.3), the relation between $P^*(t)$ and $C^*(t)$ is given by

$$C(t) = \sum_{n=-M}^{M} \alpha_n P^*(t-n) = (A^* * P^*)(t)$$

The given values α_n coincide with a given sampled function $A^*(t)$. In the frequency domain this results in

$$c^*(f) = a^*(f) \cdot p^*(f) \tag{3.9}$$

From relations (3.9) and (3.6) the values α_n may be derived by demanding

$$a^*(f) \, d_k(f) \approx 1$$

This Finite Impuls Response (FIR) design problem can be treated in different ways, see f.i. (Rabiner, L.R.'75).

<u>c</u>. Reconstruction

The relation between $C^*(t)$ and $S(t)$ is given by

$$S(t) = (C^* * B_k)(t)$$

or in the frequency domain

$$s(f) = \text{sinc}^k(f) \cdot c^*(f) \tag{3.10}$$

So, the transfer function of the reconstruction process equals the transfer function of the convolution process before sampling.
From the frequency characteristics of each of the system components the relation between $x(f)$ and $s(f)$ can easily be derived, this results in

$$s(f) = \sum_n T_n(f) \, x(f-n) \tag{3.11}$$

The functions $T_n(f)$ are for the exact deconvolution given by

$$T_n(f) = \frac{\text{sinc}^k(f) \, \text{sinc}^k(f-n)}{\sum_m \text{sinc}^{2k}(f-m)} \tag{3.12}$$

and for the local deconvolution by

$$T_n(f) = a(f) \cdot \text{sinc}^k(f) \, \text{sinc}^k(f-n)$$

$$a(f) = \sum_{j=-M}^{M} \alpha_j \, e^{-i2\pi j f}$$

herein expression (3.1) is used for $b_k(f)$.
In the sequal the local deconvolution process is not further analysed.
An interpretation of (3.11) is as follows. The frequency response $s(f)$ consists of a proper response as from a linear system ($T_o(f)$) and residual responses from a multiple displaced inputspectrum ($T_n(f)$; $n\neq o$); this latter effect will be called aliasing. If $x(f)$ is band limited to the Nyquist range, $s(f)$ will contain no aliasing into the Nyquist range, but it will contain attenuated replications of the band limited function $x(f)$, which are translated outside the Nyquist range (equation 3.11).
Assuming a more or less flat input spectrum a good measure for the aliasing effect is the function $T_A(f)$

$$T_A^2(f) = \sum_{n\neq 0} T_n^2(f) \tag{3.13}$$

$$= T_o(f)(1-T_o(f))$$

There is no aliasing where T_o is either unity or zero.
In figure A4 the response proper and aliasing measure T_A are shown for the case $k=3$ using formulae (3.12) and (3.13).

In the table below the 3 db power cut off or 70 % amplitude response is given with respect to $T_o(f)$ and some values of $T_A(f)$, for several orders k of the B-spline.
From figure A4 and table 1, one can see that $T_o(f)$ approaches the ideal zonal low pass filter with impulse response sinc(t) and this approach becomes better when the order of the spline increases. When data reduction is achieved using ideal zonal low pass filtering, all the computations involved i.e. convolution and interpolation, concern infinite series. In the approach proposed here only the deconvolution process consists of such a computation, but this has to be done once, independent of the number of interpolations. Each interpolation thereafter is easily performed using a few B-spline coefficients.

TABLE 1
Frequency characteristics of the spline approximation

k	70 % resp.	$T_A(0.5)$	$T_A(0.45)$	$T_A(0.4)$	$T_A(0.3)$
1	f= 0.324	0.491	0.499	0.494	0.440
2	f= 0.444	0.5	0.465	0.378	0.188
3	f= 0.465	0.5	0.421	0.273	0.079
4	f= 0.473	0.5	0.373	0.190	0.024
5	f= 0.479	0.5	0.323	0.129	0.017
6	f= 0.483	0.5	0.275	0.087	0.009

From table 1 one can see that aliasing decreases rapidly for increasing k, however, the computation on the convolution needs more effort for increasing k.

4. HARDWARE IMPLEMENTATION ASPECTS

4.1 Introduction

The algorithm for image processing with use of splines as described in the previous chapters leads to a data reduction and reconstruction system which processes large amounts of image data with a high input rate. Because of this high data rate, real-time processing requires special effort and therefore it is of much interest to know which part should be performed in real-time.

Because data compression is required for reduction of transmission bandwith or recording capacity on board, and because the input data stream is continuous, this should be an on-line real-time process. However, for the data reconstruction process there are three options: it can be performed on-line and in real-time, off-line and not in real-time or off-line and in real-time depending on the application.

If both real-time reduction and reconstruction are required the high data rates involved require the development of two special-purpose processors: on for on-board data reduction and another for image reconstruction on the ground.

If off-line real-time reconstruction is acceptable, the presented data reduction/reconstruction method is of special interest because the off-line real-time reconstruction of the image can be divided in a non-real-time preprocessing step (deconvolution), and a real-time replay (interpolation) step. The interpolation process lends itself well to implementation on a special-purpose processor, whereas deconvolution, which is more complicated, can be programmed on a general-purpose computer. In case of an off-line non-real-time reconstruction, both deconvolution and interpolation should be implemented on a general-purpose computer.

In this chapter, we further concentrate on the implementation of the real-time, on-board data reduction by a special-purpose processor, because this part of the processing is most critical concerning volume, power consumption etc, and is functionally identical for all applications.

All implementation structures given in this paper are proposals, so they have not been realized and tested in practice.

4.2 The algorithm

The algorithm consists of three parts as was shown in chapter 3.
Solely the first part (the correlation process) is an on-board process, yielding the decimated data-stream $Q_{v,\mu}$. Once on the ground, $Q_{v,\mu}$ is first converted into $D_{v,\mu}$ and then interpolated, which results in the reconstructed image S(s,t). In practice the image X(s,t) will be sensed with a scanner giving a representation which is continuous in scan direction and discrete in cross-scan direction. Likewise, S(s,t) is imaged with a scan-by-scan method. Therefore, $X_n(t)$ and $S_n(t)$ will be used as descriptors hereafter; n is the line number.

4.3 Implementation of one-dimensional spline correlator

4.3.1 Introduction

As indicated in 2.2 the two-dimensional spline-correlation can be constituted from two one-dimensional correlations. A simple implementation form using this property is made with two one-dimensional convolution filters with a buffering system in-between. First the information is convolved in scan-direction, giving decimated data to the line-buffering, which is needed for the second, cross-scan convolution. The line-filtering can be implemented as a continuous time filter as well as a discrete time filter, depending on the way the line-information is represented. The cross-scan filter is always a discrete time filter.

In the following sections a number of implementations of one-dimensional B-spline correlators will be outlined. Section 4.4 gives an indication of the two-dimensional filter.
In the following it is assumed that the order of the splines (k) is 3. Other orders result in similar implementations.

4.3.2 One-dimensional analog correlator

Essentially, the analog B-spline correlator can be made as a linear, time-invariant analog filter, of which the impulse response has the form of the (symmetrical) B-spline (see Fig. 4.1), followed by a sampler
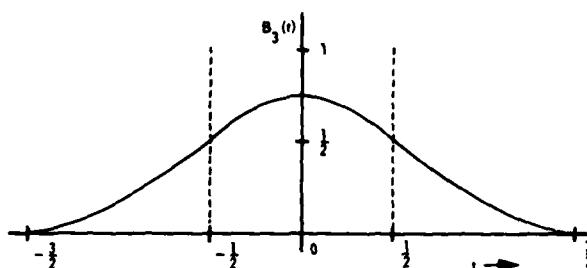
Fig. 4.1 B_3-spline function

This approach is not leading to a simple solution, because a finite-duration, symmetrical impulse response is required. However, it is possible to make use of the special properties of the B-spline correlation integral, which will lead to a hybrid solution consisting of a combined analog and digital filter. The operation to be performed is (for correlation of line information)

$$P_{n,\mu} = \int_{-\infty}^{\infty} B_3(t-\mu)\, X_n(t)\, dt$$

n = line number
t = continuous variable in line direction
μ = correlation number (integer)

(note the sample-interval scaling to 1).

The values $P_{n,\mu}$ are in fact the samples of

$$P_n(\tau) = \int_{-\infty}^{\infty} B_3(s-\tau)\, X_n(s)\, ds$$

on integer values of τ.

In order to make a design for the analog correlator first a method is found for computing $P_n(\tau)$, which is done by means of Fourier Transforms.
The Fourier transform of $P_n(\tau)$ is (using (3.1) and (3.3))

$$P_n(f) = \text{sinc}^3(f)\, X_n(f)$$

$$= [i2 \sin(\pi f)]^3 \frac{X_n(f)}{(i2\pi f)^3}$$

The function $X_n(f)/(i2\pi f)^3$ is the Fourier transform of the third order primitive of $X_n(t)$, denoted by $X_n^{[-3]}(t)$. The function $i2 \sin(\pi f)$ is the Fourier transform of $\delta(t+\tfrac{1}{2}) - \delta(t-\tfrac{1}{2})$. Using the convolution theorem (3.3) we can express $P_n(\tau)$ in terms of $X^{[-3]}(\tau)$:

$$P_n(\tau) = (X_n^{[-3]}(s) * [\delta(s+\tfrac{1}{2}) - \delta(s-\tfrac{1}{2})]^{*3})(\tau)$$

$$= X_n^{[-3]}(\tau+3/2) - 3X_n^{[-3]}(\tau+\tfrac{1}{2}) + 3X_n^{[-3]}(\tau-\tfrac{1}{2}) - X_n^{[-3]}(\tau-3/2)$$

Thus the samples $P_{n,\mu}$ of $P_n(\tau)$ can be written as

$$P_{n,\mu} = X_n^{[-3]}(\mu+3/2) - 3X_n^{[-3]}(\mu+\tfrac{1}{2}) + 3X_n^{[-3]}(\mu-\tfrac{1}{2}) - X_n^{[-3]}(\mu-3/2)$$

Accepting a time delay of $\frac{3}{2}$, we obtain the expression

$$P_{n,\mu+\frac{3}{2}} = X_n^{[-3]}(\mu+3) - 3X_n^{[-3]}(\mu+2) + 3X_n^{[-3]}(\mu+1) - X_n^{[-3]}(\mu)$$

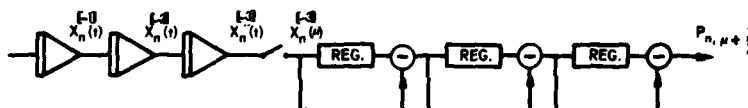This leads to the realisation of figure 4.2.

Fig. 4.2 Structure analog B-spline correlator

This realisation is built-up with a series of three integrators, serving as an analog presampling filter, a sampler, and a very simple digital filtering operation.
A drawback of this realisation is the fact that the integrators will show overflow after some time. Because the information is processed line after line the resulting harm may be kept within bounds but anyway the large dynamic range needed for the integrators and thus for the analog-to-digital conversion, is still a disadvantage. Currently, it is investigated at NLR whether a compensation for the overflow by means of a controlled input offset is possible.

The processing capacity of this realization, that is the maximal number of correlations that can be computed, is determined by the subtraction time. In case three parallel subtractions are used with a cycle time of 100 ns, approx. 10^7 correlations per second are computed.

When the dynamic-range problems of the integrators are set aside, this hybrid solution is attractive because of its relative simplicity and the large processing capacity. Moreover the delivered correlation coefficients $P_{n,\mu}$ are basically exact, as opposed to the output of the full digital realizations, as described hereunder.

4.3.3 One-dimensional digital correlator

The B-spline correlator may also be realised as a discrete time filter with a sample stream as input. In that case the real correlation in scan direction is approximated by:

sampling : $X_{n,m} = X_u(\frac{m}{v})$

discrete correlation: $P_{n,\mu} = \sum_k B_3(\frac{k}{r} - \mu) \, X_n(\frac{k}{r})$

This approximation is attained by sampling the input signal with sampling frequency r, and correlating it with a sequence obtained by sampling $B_3(t)$ frequency r. In this way an FIR-filter (finite-duration impulse response filter) is generated, serving as a decimation filter with reduction factor r. The sampling frequency must be sufficiently high in order to limit aliasing in input signal as well as in the B-spline. It is possible to avoid aliasing in the input signal by inserting a presampling filter, but this will produce additional linear distortion. In practice the aliasing error in the representation of the B-spline as a series of samples is acceptable for $r \geq 4$.

In cross-track direction there is no choice between continuous and discrete correlation: because of the line-scanning the information to be correlated is always discrete. The corresponding operation is here:

$$Q_{v,\mu} = \sum_k B_3(\frac{k}{r} - v) \, P(\frac{k}{r},\mu)$$

$$\text{where } P(\frac{k}{r},\mu) = P_{k,\mu}$$

The discrete implementation methods described here are worked out for a reduction factor (r) equal to 4. Then, the operation to be implemented takes the form:

$$y_{4p} = \sum_{i=0}^{11} a_i x_{4p-i} \quad , p \text{ is an integer.}$$

This corresponds to a vector multiplication, of a fixed coefficients vector with a variable sample vector, built-up with 12 consecutive samples.

The 12 coefficients a_i are listed in table I. An interesting fact is that the binary representation as a fixed point fraction needs only 7 bits, without rounding.

Further, it is assumed that the input samples are quantized in 8 bits.

TABLE I

The correlation coefficients a_i, for third order spline (k= 3) and reduction factor r = 4.

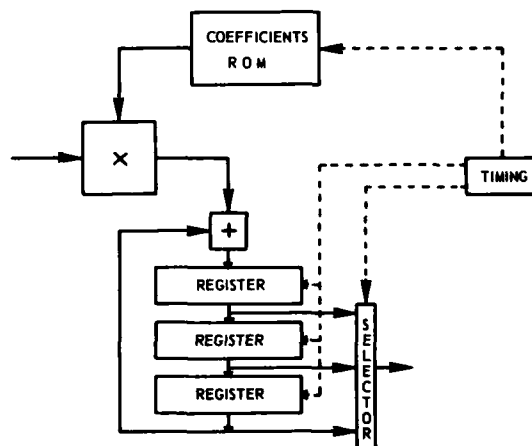| i	a_i(dec)	a_i(bin)
0	0,001953125	0000001
1	0,017578125	0001001
2	0,048828125	0011001
3	0,095703125	0110001
4	0,15234375	1001110
5	0,18359375	1011110
6	0,18359375	1011110
7	0,15234375	1001110
8	0,095703125	0110001
9	0,048828125	0011001
10	0,017578125	0001001
11	0,001953125	0000001

Fig. 4.3 Structure digital B-spline correlator with multiplier and coefficients ROM

4.3.3.1 Method 1. Application of multiplier and coefficient ROM

An obvious implementation method uses a hardware multiplier and a read-only-memory (ROM) for storing the coefficients a_i. A proper sturcture (see fig. 4.3) is obtained by immediately calculating all relevant contributions to the vector product series, for every new input sample; so, for k= 3, there are 3 multiplications and 3 additions at a time. The vector products arise as accumulations of contributions from 12 input samples each.

In this way, only three intermediate results have to be stored.

This is of special interest for the processing in cross-scan direction, see 4.4.1.

The coefficients ROM is just small, 12 x 7 bits. The adder and the registers must be at least 17 bits wide, if full length calculations are required without round-off error. In that case, a processing capacity of approx. 550.000 correlations/s, with a cycle time of 150 ns, is easily attainable with commercially available components.

The structure of this implementation depends on the reduction factor required. Note, however, that the processing capacity, in correlations/s, does not depend on this reduction factor.

4.3.3.2 Method 2. Application of special vector multiplier

Because the operation to be performed is a vector multiplication of a variable (sample) vector with a fixed (coefficients) vector, the vector multiplier as first described by Peled and Liu, and Little (Peled, A,'74 and Little, W.D.,'74) can be applied. Its principles are as follows.

$$\text{The operation} \quad : \quad y_n = \sum_{i=0}^{M-1} a_i x_{n-i}$$

$$\text{may be written as:} \quad y_n = \sum_{i=0}^{M-1} a_i \sum_{j=1}^{N} x_{n-i,j} 2^{-j}$$

when the x_{n-i} are N-bits positive fractions.

Definition of $C_{n,j}$ and interchanging of the order of the summations give

$$y_n = \sum_{j=1}^{N} C_{n,j} 2^{-j}$$

$$C_{n,j} = \sum_{i=0}^{M-1} a_i x_{n-i,j}$$

The coefficients $C_{n,j}$ can assume just 2^M different values, which can be calculated in advance and stored in a ROM. Instead of being calculated each time, they now can be selected by means of the set $x_{n-i,j}$ which are the jth bits of the M-values x_{n-i}. Next, the selected coefficients $C_{n,j}$ are summed in a shifting accumulator, i.e. the intermediate results are shifted over 1 bit after each addition. Figure 4.4 gives the structure of this method.

A disadvantage of this method may be the large ROM needed for the coefficients $C_{n,j}$. In case a high processing speed is required, it must be built-up with bipolar devices. An alternative structure is obtained by making use of the symmetry in the series a_i: $a_0 = a_{11}$, $a_1 = a_{10}$ and so on. The operation:

$$y_n = \sum a_i x_{n-i}$$

can therefore be written as:

$$y_n = \sum_{I} a_i x_{n-i} + \sum_{II} a_i x_{n-i}$$

where the sets a_i are identical for both I and II. The summations I and II can then be calculated separately, making use of the same ROM; its size is now reduced to 64 x 9 bits. Finally both partial vector products must be summed.

In order to adapt the vector-multiplier structure to the specific task of fast spline correlation it should be modified such that it can cope with a continuously running input data stream. This is accomplished by inserting registers between the arithmetic elements, which results in a pipeline structure, and by modifying the input register-file structure, so that the computation is repeated every time four new samples are received.

The processing speed attained with 100 ns cycle time is approx. 1.2 million correlations per second. The speed is halved in the alternative structure, with ROM and accumulator multiplexed.

The structure, again, depends on the reduction factor, in contrast with the upper bound for the number of correlations per second.

4.3.3.3 Method 3. Application of accumulators, based on properties of the B-spline

In contrast with both previous methods, the accumulator method described here is limited to application to B-spline correlation. The method has a strong relationship with the integrator method for analog correlation (see 5.3.1).

The operation to be performed is:

$$y_{4p} = \sum_{i=0}^{11} a_i x_{4p-i} \quad , \quad p \text{ integer.}$$

When we overlook for a while that just one to four correlations have to be calculated, application of the z-transform yields:

$$Y(z) = X(z) \sum a_i z^{-i} \quad .$$

Hence, the transfer function $H(z) = \sum a_i z^{-i} =$

$$= \frac{1}{512} (1 + 9z^{-1} + 25z^{-2} + 49z^{-3} + 78z^{-4} + 94z^{-5} + 94z^{-6} + 78z^{-7} + 49z^{-8} + 25z^{-9} + 9z^{-10} + z^{-11}).$$

Factorization yields:

$$H(z) = \frac{1}{512} (1 + z^{-1} + z^{-2} + z^{-3})^3 (1 + 6z^{-1} + z^{-2})$$

$$= \frac{1}{512} (\frac{1-z^{-4}}{1-z^{-1}})^3 (1 + 6z^{-1} + z^{-2}).$$

This expression lends itself to the simple implementation as shown in figure 4.5, which consists of mere additions, subtractions and registers, and one multiplication with a factor 6.

This multiplication is, however, easily accomplished with one additional adder.

Apparently, there could be problems due to overflowing accumulators, but this is not the case because modulo 2^{17} arithmetic is applied. This is possible because the range of the resulting values is known. The range determines the number of bits involved in an addition or subtraction, which influences the end result. Therefore, overflow in the intermediate results is of no importance. When 8 bits samples are used, all computations are done in 17 bits.

For a practical implementation according to this method.it is taken into account that just one to four correlations have to be calculated. Hence a number of additions, subtractions and registers in the final part of the structure can be deleted, resulting in a total of 5 adders/subtractors and 8 registers, and addition of a simple control structure.

The processing speed, attainable with commercially available hardware is approx. 2.5 million correlations/s, with a 100 ns cycle time. This number decreases with higher reduction factors.

4.4 Implementation of two-dimensional spline correlator

4.4.1 Introduction

The two-dimensional spline correlation consists of two one-dimensional correlators (Fig. 4.6), of which the first works in scan direction, and the second in cross-scan direction. The first correlator processes the line information, line-after-line; the second one operates simultaneously on n_1 correlations (n_1 is reduced number of pixels or a line). This parallelism necessitates a form of intermediate buffering.

The size of the intermediate buffering depends on the implementation method chosen for the second correlator; further on the reduced number of pixels per line. The relation is:

$$M = \frac{n_1}{r_1} \times R.$$

M = buffer size in words
n_1 = number of pixels/line
r_1 = reduction factor in line direction
R = number of registers in one-dimensional implementation.

The choice for the second correlator is therefore mainly determined by the number of registers. There is no relation between the size of the intermediate buffer and the chosen implementation method for the first correlator; hence this choice depends on more general requirements.

4.4.2 Analog correlation in line direction and digital correlation in cross direction

If represented as an analog signal the information in line direction can be correlated by the analog method. The output of this line filter is a sequence of numbers for every line. Next, the sequences are rearranged in order to be processed in cross-scan direction, in the second correlator.

For the second correlator, the digital implementation is chosen with a minimum number of registers, resulting in method 1, section 4.3.3.1.

Assuming that only the digital correlation determines the upper bound for the processing speed that can be obtained, the maximum output data rate is 550.10^3 pixels/s.

4.4.3 Digital correlation in both line and cross directions

If the image information to be processed has been sampled before, or if full digital processing is preferred, both correlators are digital. For the second correlator (cross direction) there is no difference as compared with the previous section. The choice for the first correlator, however, is not mainly determined by the number of registers, but for instance by the complexity and amount of arithmetic elements.

For the first correlator mainly the methods 1 and 2 can be chosen. Method 2 is only considered if the large scale integrated multiplier-I.C's can not be applied, for instance because of the high expense.

Method 1 is especially attractive in case of moderate requirements on the processing speed, so that the first correlator can be partly combined with the second one, which is also implemented according to method 1. Especially the arithmetic elements can be shared.

The highest speed is attainable with method 3. The disadvantage of the somewhat higher component count is alleviated by the fact that mere low-cost standard TTL-components are used.
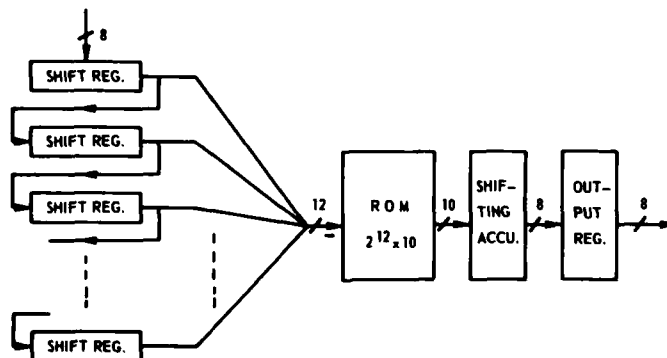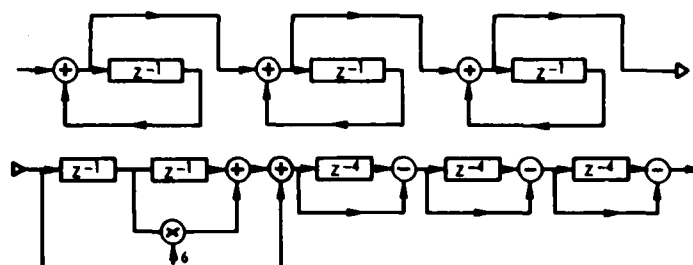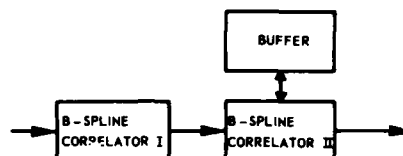

Fig. 4.4 Structure ROM vector multiplier

Fig. 4.5 Structure accumulator method for B₃-spline correlation

Fig. 4.6 Structure two-dimensional B-spline correlator

5. EDGE-ENHANCEMENT

From the results (see chapter 6) obtained with the image approximation method as outlined before, one can see that sharp edges in the original image are smeared out in the image approximation. For some purposes this may be an undesired phenomenon. In order to suppress this smearing effect, an edge enhancement algorithm is designed. The idea of the algorithm is described with help of the figure below. It is assumed that the measured radiance from the sensor is quantized in 8 bits. The radiance levels, 0-255, are divided now in 2^N intervals (N is a fixed number between 0 and 8). In figure 5.1, 4 intervals are distinguished with equal size. However, these sizes may also be taken in an adaptive way.
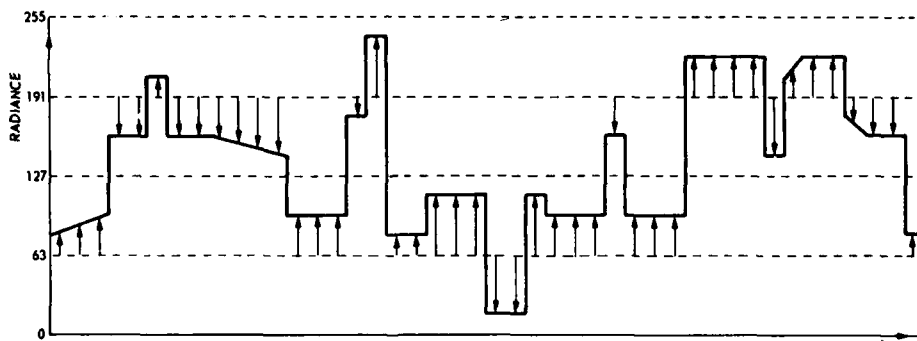

Fig. 5.1 Changing to relative radiance within an interval

In applying the algorithm on the input signal $X(n,m)$, two data streams are created. Stream 1 contains the information about the interval for each pixel, the second stream contains the relative radiance level of each pixel within the interval, as is showed in figure 5.1 for one scanline. For the example in figure 5.1 the two data streams are plotted explicitly in the figures 5.2 and 5.3.

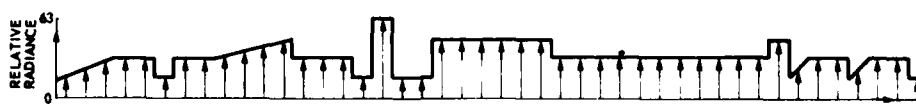
Fig. 5.2 Stream 1, after edge filtering

Fig. 5.3 Stream 2, after edge filtering

Data stream 2 is in fact a nonlinear conversion of the input signal, as follows:
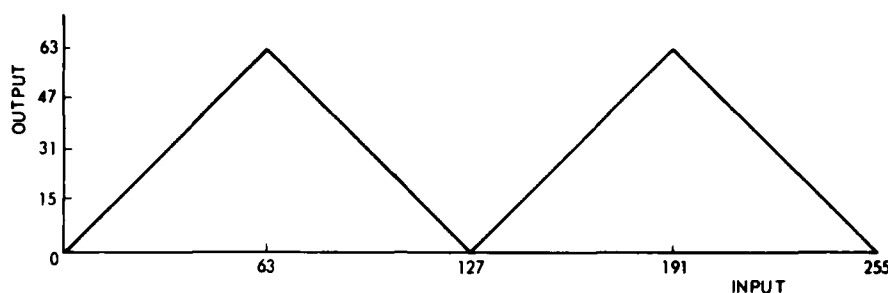
Fig. 5.4 Stream 2 versus input to edge filter

It is clear that it is possible to reconstruct the image X(n,m) from stream 1 and stream 2, with a so-called inverse edge filter. After applying the edge filter, stream 1 stays unaltered but stream 2 is seen as a new image (it has less edges than the image X(n,m)) and is approximated with a spline S(n,m) as outlined in chapter 2. Finally, stream 1 and the approximation S(n,m) are added in the inverse edge filter, giving the approximation $\hat{S}$(n,m) to the image X(n,m).
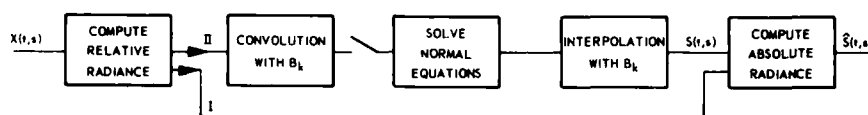Resuming, the edge filter is used in the following system configuration.

Fig. 5.5 Image approximation with edge-enhancement

By taking N big enough (determining the number of intervals) the accuracy of $\hat{S}$(n,m) becomes better, but data stream 2 consists of more bits describing the intervals. However, for small N these bits are very redundant and therefor can be encoded such that less bits are required for transmission c.q. storage of stream 1.

6. SIMULATION RESULTS FOR LANDSAT DATA

6.1 Preliminary results of approximation with splines for LANDSAT data

The simulation results described in this chapter have been obtained by combining computer programs to process part of a LANDSAT Computer Compatible Tape (CCT) of real remote sensing data. The area chosen is around the town of Harderwijk, containing remnants of Lake IJssel, regular linear structures in the newly reclaimed Lake IJssel polders and irregular fields of a rich structure.
Consequently high spatial frequencies are present in the data. The dimensions are 240 x 240 pixels. The results shown (photo 1-3) are of the original data, a least-squares parabolic spline approximation and an approximation obtained with averaging 2 x 3 pixels. The spline intervals are also 3 pixels in horizontal lines and 2 pixels in vertical lines. Each surface is neglecting boundary effects, described by 80 x 120 coefficients for each colour (resulting in a data-reduction of a factor six).
Computing the approximating image required about 27.5 CPU seconds on the CYBER 72 at NLR. About 2.5 CPU seconds were needed for solving the normal equations (deconvolution) using the Cholesky method and 12.5 seconds for both the computation of the B-spline correlation data and the reconstruction of the image. As was shown in chapter 4, a large gain in processing speed is attained with a special purpose processor for the correlation process. The same result is expected for the deconvolution and interpolation processes.
Comparing photo 1 and photo 2, it is clear that a number of features in the original image have been smeared out or reappear at displaced positions in the reconstructed image. This was expected due to the Low pass filter characteristics of this approximation method resulting in a decrease of spatial resolution. However, down to a fair level of detail the overall structure is reproduced better than the approximation obtained with averaging 2 x 3 pixels (photo 3).
One can see that especially edges from the original image are smeared out in the approximation. Therefore, in chapter 5 an edge enhancement algorithm was designed in order to have a better edge reproduction in the approximation. In the next section some results are given after application of this edge enhancement algorithm in combination with the least-squares approximation method.

6.2 Preliminary results of approximation with splines and edge enhancement for Landsat data

Using the system configuration of figure 5.5, a B-spline approximation with edge enhancement as outlined before, is applied on an image of the area around the town of Harderwijk. The result is shown on photo 4. The value of 2^N, denoting the number of intervals in the radiance levels is chosen to be 4 for the green colour and 8 for the red colour. The four intervals are bounded by the points (i-1)x 64-$\frac{1}{2}$, i=1(1) 5 and the eight intervals are bounded by the points (i-1)x 32-$\frac{1}{2}$, i=1(1) 9.
The difference in intervals is due to the fact that in the red colour the most significant bit in the original image is not varying.
The choice of the boundary points is independent from the image to be approximated, so better results may be expected when these points are chosen in an adaptive way.
The spline intervals are 3 pixels in horizontal lines and 2 pixels in vertical lines. The overhead information from the edge filter (stream 1) is about 1 bit/pixel for each colour. Thus, the total data reduction achieved is a factor 4
Comparing photo 1 and photo 4 one can see that a big part of the edges present in the original image are enhanced in the approximation. However, there are also vague areas in the approximation, it seems that these areas are copies from the approximation using only B-splines (photo 2). The reason for their appearance is

that the number N was to low for their detection or/and the boundary points are not taken optimal. Another remarkable phenomenon is the occurrance of isolated dark and light coloured pixels photo 4 only. This is a result of application of the edge enhancement algorithm when an edge in the original image is not precisely divided over the two intervals which are defined by the boundary points. However, when the boundary points are taken optimal this phenomenon will be less worse.

7. CONCLUSIONS

1. Least-squares image approximation using splines is a good method for real-time data reduction on-board, followed by an image-reconstruction on the ground.

2. The method is a special case of low-pass fittering and therefore results in a decrease of spatial resolution. When the order of the spline increases, the method converges to the ideal zonal low-pass filter and the aliasing effects disappear.

3. For the data reduction step it is possible to develop a special-purpose processor for real-time application up to an output data rate of 550.10^3 pixels/s.

4. The main computational effort in the data processing on the ground lies in the preprocessing of the compressed data. The image interpolation from the modified compressed data is relatively simple.

5. The method is especially tuned to applications with simple data reduction hardware on-board and a more extensive data processing on the ground.

6. Real-time reconstruction of the image is difficult due to the complex preprocessing of the compressed data. However, by simplifying or deleting the preprocessing, it is possible to have a Quick-look facility offering lower image quality.

7. The proposed edge-enhancement algorithm, which adds some extra information to the compressed data set, can improve the image quality considerably.

ACKNOWLEDGEMENT

REFERENCES

1 De Boor, C., A practical guide to splines. Appl. Math. Series 27, Springer Verlag 1978.

2 Little, W.D., An algorithm for high speed digital filters. IEEE Trans. Computers, Vol. C-23, pp. 466-469, 1974.

3 Peled, A. and Liu, B., A new hardware realisation of digital filters. IEEE Trans. Acoust. Sppech, Signal Proc., Vol. ASSP-22, pp 456-462, 1974.

4 Raliner, L.R. and Gold, B., Theory and applications of digital signal processing. Prentice Hall, 1975.

5 Renes, J.J. and Beintema, D., The use of B-spline approximation and of an array processor in the IRAS ground operations and preliminary analysis facility, in: R. Holdaway (ed.), Proceedings of 3rd B.I.S. computers and space technology conference on image processing techniques applied to astronomy and space research, Appleton (U.K.), November 1979.

APPENDIX A, SOLVING THE NORMAL EQUATIONS

In approximating the input function $X(t)$ on the interval $[0,T]$ with a spline $S(t)$ of order k in the least squares sense, one has to solve the normal equations (2.6). In vector notation, one has to solve the linear problem.

$$A\underline{c} = \underline{p} \tag{A.1}$$

The square matrix A consists of zeros except on the main diagonal and the 2k-2 upper- and lower diagonals. The vector $\underline{c}$ denotes the unknown B-spline coefficients $\{C_\nu\}$ and the known vector $\underline{p}$ denotes the left part of equation (2.6).

For the case k=3, the matrix A takes the form

$$A = 1/120 \begin{bmatrix} 6 & 13 & 1 & & & & & & & & \\ 13 & 60 & 26 & 1 & & & & \emptyset & & & \\ 1 & 26 & 66 & 26 & 1 & & & & & & \\ & 1 & 26 & 66 & 26 & 1 & & & & & \\ & & 1 & 26 & 66 & 26 & 1 & & & & \\ & & & & & & & & & & \\ & & \emptyset & & & & 1 & 26 & 66 & 26 & 1 \\ & & & & & & & 1 & 26 & 60 & 13 \\ & & & & & & & & 1 & 13 & 6 \end{bmatrix} \tag{A.2}$$

Finding the exact solution $\underline{c}$ of (A.1) is a time consuming effort, however, for some applications an approximation of the exact solution $\underline{c}$ may be acceptable, f.i. for use in a Quick Look facility. For real time applications a compromise has to be found between the quality of the approximation and the time needed for computing the approximation.

One way of approximating the solution to equation (A.1) is computing an approximation $\{C_\nu\}$ of the form

$$C_\nu^* = \sum_{m=-M}^{M} \alpha_m P_{\nu-m} \qquad (A.3)$$

In this approximation one is inspired by the fact that the coefficients b_{ij} of A^{-1} satisfy the decay rate bound.

$$b_{ij} \le \text{const.} \quad \lambda^{|i-j|}$$
$$|\lambda| < 1$$

The value of λ depends upon the order k of the B-spline. In chapter 3 this type of approximation is also treated in the frequency domain.

In finding the almost exact solution of (A.1) an efficient method is found faster than the approximation (A.3). A disadvantage of this method will be discussed later. The method is based upon a kind of Cholesky decomposition of the matrix A, i.e. one can write the matrix A of (A.1) as

$$A = C \, D \, C^T \qquad (A.4)$$

only for the last (k-1) rows of A this relation is not valid, but this is neglected in the beginning, a correction for this is made later on.
The matrix C is an upper triangular matrix with non zero elements on the main- and (k-1) upper diagonals, moreover the matrix elements on the same diagonal are equal.
The matrix D is the Chronecker sum of a non-singular (k-1) x (k-1) matrix and the identity matrix.
Thus applying the decomposition (A.4) one can easily solve (A.1) using recurrent relation yielding the solution $\hat{c}$. However, the composition (A.4) is not unique, it should be chosen such that the linear system $Cx = b$ can be solved in a stable manner using recurrent relations. For k=3, the decomposition of A from equation (A.2) is given below.

$$C = \begin{bmatrix} 7.3409 & 3.4772 & 0.1362 & & & \emptyset \\ & 7.3409 & 3.4772 & 0.1362 & & \\ & & 7.3409 & 3.4772 & 0.1362 & \\ & \emptyset & & 7.3409 & 3.4772 & 0.1362 \\ & & & & 7.3409 & 3.4772 \\ & & & & & 7.3409 \end{bmatrix}, D = 1/120 \begin{bmatrix} 0.0901 & -0.1883 & & \\ -0.1885 & 0.8887 & \emptyset & \\ & & 1 & \\ & & & 1 \end{bmatrix}$$

Finally the necessary correction should be made on the solution $\hat{c}$ due to inaccurate decomposition (A.4). From the discussion above it follows that the effect of an element p_n from the input data $\underline{p}$ dies out exponentially in the solution sequence $\{C_\nu\}$ (controlled by the matrix C). By demanding a prescribed accuracy in the solution $\{C_\nu\}$ one can determine the accurate part of the solution $\hat{c}$. The inaccurate part of $\hat{c}$ should be computed again with use of the decomposition:

$$A = C^T \, D' \, C$$

where the matrix D' in the Chronecker sum of the identity matrix and the previous non-singular (k-1) x (k-1) matrix.

A disadvantage of the decomposition method is that the input data $\underline{p}$ has to be passed two times in opposite directions.
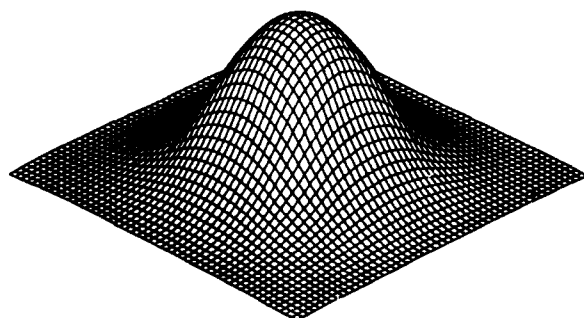

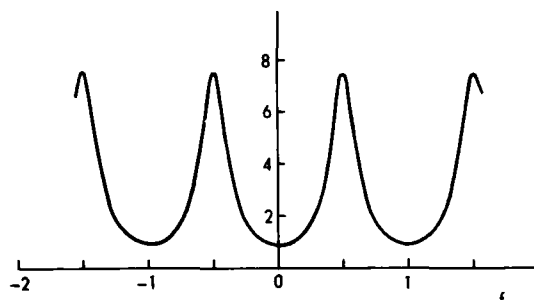Fig A.1 Iso-metric projection plot of parabolic B-spline

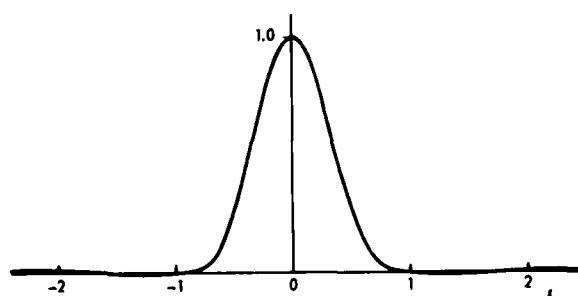Fig. A.3 Plot of the function $(1-\sin^2(\pi f) \ 2/15 \ \sin^4(\pi f))^{-1}$

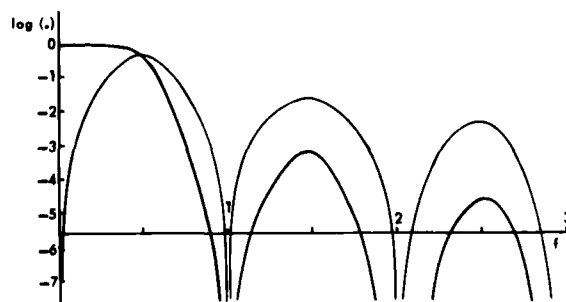Fig. A.2 Graph of the function $\text{sinc}^3(f)$

Fig. A.4 Graph of logarithms of the proper transfer function $T_0(f)$(fat) and the aliasing measure $T_n(f)$(thin)

14-14

Photo 1: original

Photo 3: after averaging 2x3 pixels

Photo 2: after third order spline approximation

Photo 4: after third order spline approximation with edge enhancement

A FLEXIBLE IMAGE PROCESSING SYSTEM

Peter Gemmar, Hartmut Ischen, Karl Lütjen
Forschungsinstitut für Informationsverarbeitung
und Mustererkennung (FIM/FGAN)
Breslauer Straße 48, 7500 Karlsruhe
Fed. Rep. Germany

SUMMARY

A multiprocessor system for image processing is presented. Structural requirements are shown for frequently used image processing tasks. Design principles and implementation techniques to realize high performance capabilities in a multiprocessor system are described in terms of simultaneous processing, structural flexibility, and data input/output. The FLIP (Flexible Image Processing system) is operated in conjunction with a host computer and can perform up to 64 MIPS (Million Instructions Per Second). FLIP can be considered as a multi-pipeline-processor comprising 16 individual processors and a high speed data input/output processor. The structural flexibility of FLIP is achieved by its individual processors. They can be arranged by programming in nearly any manner to adapt the hardware to an optimal processing structure of the function to be processed. FLIP achieves simultaneous processing by combining parallel processing and pipelining of operations. Since sychronization of the individual processors is data-flow controlled the structural behaviour of the processing system is easy to survey. A special data exchange processor provides convenient and rapid access to image data especially for all kind of homogeneously performed window operations.

Some practical applications (e.g. image differentiation, image convolution) are explained to demonstrate FLIP's system performance. The experimental results of the developed system show, that FLIP reduces image processing times by factors between 10 to 100 compared with conventional techniques (e.g. using a 1 MIPS general purpose computer).

1. INTRODUCTION

Processing of pictorial data on a conventional sequential computer is very time consuming. This is not only a trouble for real-time applications, as it is in most military applications, but also in the field of experimental simulations. Picture processing requirements have increased so much in the past, that it is unrealistic to believe that today and in the near future the desired processing times can be met by conventional computers. To overcome this problem many special image processing computers with parallel processing capabilities have been proposed.The performance of such systems is often evaluated by comparing the execution time of a specific image processing task on the system under consideration versus the execution time of the same task on a sequential computer.

The architecture of image processing systems is mostly driven by the structural requirements of image processing tasks. The evaluation of those requirements show that for the various steps of image processing different structures are needed.

With the Flexible Image Processing System (FLIP) we designed and built a flexible processing structure which combines simultaneous processing and structural flexibility. Its main application is computing of homogeneous and parallel operations often used for picture preprocessing and feature extraction. Those parallel operations are spatial image filtering, template matching, local correlation, ect.

1.1. Structural Requirements for an Image Processing System

To determine the processing structure of a new system it is nessessary to evaluate the structural requirements of typical image processing tasks. The following simple algorithm, called "stroke difference" is an example.

With the submatrix notation of Figure 1 the stroke difference is given by

$$Gx(x,y) = |(F(1,1) + F(2,1) + F(3,1) - (F(1,3) + F(2,3) + F(3,3))|/3$$

$$Gy(x,y) = |(F(1,1) + F(1,2) + F(1,3) - (F(3,1) + F(3,2) + F(3,3))|/3$$

$$G(x,y) = (Gx(x,y) + Gy(x,y))/2$$

This processing is a local or window operation. The picture elements are taken from a submatrix (window) which is systematically shifted over the whole image and the processing is repeated without change (i.e. homogenously) for each window position.

To evaluate the structural requirements it is useful to draw the corresponding computing graph of the mathematical expression. For this purpose the algorithm is decomposed in elementary operations. For the above example the operations are displayed as nodes in the processing flow by connecting arrows in Figure 2.

The computing graph shows two specific principles:

- Parallel processing: This type of processing is characterized by the concurrent computation of unrelated operations (e.g. operations 1-4 in the first stage (column), 5-8 in the second stage, ect.).

- Pipelining: This type of processing is characterized by the sequential data flow through several processors (if we use for each operation one processor) and the parallel working of all stages (e.g. the operations 1, 5, 9, 11, 13 and 14 constitute a pipeline).

Both principles are widely used to speed up computation. The computing graph of our example shows the combination of these two principles, what briefly can be described as simultaneous processing. The speed of computation is dictated in this structure by the slowest operation within one pipeline. The problem of system optimization is the distribution of operations in such a way, that each processor uses nearly the same computation time.

The design of the FLIP was greatly influenced by the evaluation of image processing tasks like the above mentioned stroke difference method. As a result, the FLIP comprises the following general features:

- Multiprocessor system: High computation power is provided by 16 processors

- Structural programmability: It is possible to arrange the processors in any desired structure by software means

- Parallel input data stream: A programmable data exchange processor (PEP) provides a high capacity parallel input data stream

- Effective and simple synchronization: The synchronization of the 16 processors is data flow controlled and supported by hardware

- Programmable operations: Each processor can be programmed to perform complex and different operations.

2.   IMAGE PROCESSING SYSTEM FLIP

2.1.   FLIP System Configuration

The FLIP system is connected to a host computer (Figure 3) and is operated like a peripheral device. Image data to be processed often comes from a disk and after processing the results are also stored on a disk or alternatively they are displayed on a raster-display. The host computer itself is part of a larger system and is also connected to two other computer systems (with additional image input/output devices). So, the processing capabilities of the FLIP are also available to users of the two other systems.

The FLIP consits of the central processing unit FIP (Flexible Individual Processors) and the data exchange processor PEP (Peripheral data Exchange Processor). The FIP itself has 16 IP's (Individual Processors) and the PEP, which is mainly a fast and intelligent buffer device, has a fast bipolar memory of 24 KBytes and three internal processors. Additionaly the FLIP is connected to a MOS image memory (768 KBytes) with a high speed data path.

Image processing with the FLIP is usually done on a data stream flowing from the disk to the PEP, from the PEP to the FIP, and after processing within FIP, via PEP back to the disk. The three data streams, input, processing and output are flowing in parallel. Figure 4 shows this data stream from the viewpoint of image processing. Out of the input image, which is transfered pixel by pixel (pixel = picutre element), the PEP forms a parallel data stream for processing. For this, the PEP holds at least those data, which are simultaneously required by the FIP, e.g. the lines of the image covered by the window currently being processed. The FIP merges the parallel data stream back to a single data stream, which is subsequently stored as the output image.

This double funnel-shaped data flow is a feature often found in image processing tasks. Our example of the "stroke difference" operates exactly in this manner.

2.2.   Processing Unit FIP

The realized FIP-System with 16 individual processors pocesses all the capabilities which are required to achieve a flexible and programmable structure. The system works without any external control for synchronization and the data flow is used to synchronize the individual processors. A processor excuting an input instruction will wait until valid data are present on the requesting port. An extensive bus system is used to physically connect all individual processor with each other (Figure 5). Each individual processor has its own output bus with eight parallel bits of data. Each of these buses is connected to the input ports of all other processors. To increase the system speed and to ease the programming of desired structures each processor is equipped with two independent input ports. By this, each individual processor may have three data transfers at the same time, two at its input ports and one at the output bus. The instruction format reflects this feature by the provision of three adress fields, so each individual processor is a three adress machine.

All individual processors of FIP are identical. In order to achieve high processing speed, data and instructions are internally separated both in storage and signal path. In each processor the storage capacity is 50 bytes (8 bit) for data and 256 words for instructions (32 bit). Figure 6 shows the functional block diagram of an individual processor. Two independent input control units (IAC, IBC) can simultaneously select one bus out of 16 buses. The output control (OC) has its own one byte buffer allowing output data transfer to be done in parallel with program excution.The internal program execution is controlled by an asynchroneous network. Internal data are represented by eight bits. The instruction repertoire of each IP includes the basic arithmetical and logical instructions, the 8 bit by 8 bit multiplication, shift instructions, and instructions to handle multiple precision arithmetic. Subroutines can be called and special control instructions (e.g. stop instruction) are implemented. The execution time of each instruction is dependent on the complexity of its operation and adressing mode. The measured mean execution time is 250 nsec.

The data input to FIP is performed by the PEP on additional data buses. To achieve a parallel data stream (as required by the chosen structure of an image processing task) the FIP and PEP are connected with 16 independent buses (buses A and B of Figure 5) providing a possible data rate of 45 Mbyte/s. Unlike the FIP internal buses the A and B buses carry destination addresses and the PEP can chose any of the 16 IPs as destination. For data output from the FIP three buses are provided which are connected to the PEP, the MOS image memory and the host computer (with DMA), respectively. The data input/output is done in parallel with program execution as well as the FIP internal data flow on the FIP-buses. An additional control and instruction bus directly controlled by the host computer is provided for program loading and control purposes.

2.3.   Data Exchange Unit PEP

The data flow between FLIP and the host computer is accomplished by three exchange processors (EP), a data memory of 24 KB, and associated interfaces, respectively, all comprised within PEP (Figure 7). PEP is designed to provide input data for the FIP processing unit and mainly to support the homogeneous type of image processing. A multiple data stream is formed from the single input data stream coming from the host system. Working on a 10 x 10 submatrix, for example, the data rate to FIP can be 100 times greater than the rate of the sequential input data stream.

The access to the data stored in the data memory is free programmable and almost unlimited. By this, a flexible use of the FLIP system is guaranteed. Additional addressing hardware is available for the homogeneous type of image processing, so that the user must not concern about the special problems for this type of data in- and output. For window operations, the size of the data memory (24KB) limits the maximum image line lengths to be directly processed. For example, with a line length of 240 we can access a submatrix of 100 x 100 pixels window size, whereas with a line length of 8000 we can only handle a window of 3 x 3 pixels. Since FLIP acts like a peripheral device it must be controlled and supervised by the host computer. This means that within the host a program must provide at least two independent data streams:

- One input data stream usually originating from a mass storage device (e.g. disk).

- One output data stream going to a mass storage device or a display device.

Within the FLIP system the input data stream is handled by the input/output controller (Figure 7) and transferred to the PEP data memory. The multiple data stream to the FLIP is controlled by the three PEP processors. The data then are temporarely stored in the data buffers and transferred on 16 buses to the FIP processors. The results from the FIP are returned and transferred to the host or to other storage devices under the control of the input/output controller.

2.4.   FLIP - Principle of Operation

To establish a desired processing structure on FLIP, the individual processors of FIP as well as the PEP-processors (EP1-EP3) must be programmed. As soon as all programs are loaded the processing structure is latent in the system and processing can be started. This is usually intialized by the data transfer from the PEP to the FLIP. The PEP itself gets the data from the host computer by DMA (Direct Memory Access). This data flow scheme provides the host computer with complete control over the execution of the entire processing. Within FIP any desired structure can be established by programming from the pipeline over the cascade up to other parallel structures. Two simple examples are illustrated in Figure 8. The only means to establish a connection between two FIP-processors is by the address parts (operand address = bus address) within the instruction words of the processor requesting the data. In this way, the construction of a logical data path is done by the processors receiving the data. In that processors which are to send the data, output-instructions according to the data requests must be programmed (see example in next section). As soon as valid data is present on a selected bus it is latched and confirmed by the input port of the processor executing the input-instruction. Since each processor possesses two independent input ports the temporal appearance of data from different buses is of no importance. In the case of an input operation without valid data at one or both input ports, respectively, the processor will remain in a waiting state.

The data output is initiated after the execution of an instruction containing the addressing mode 'data to output-bus'. The data is subsequently latched for output and program execution is resumed. If the last data wasn't taken by another processor the execution of a second instruction with output to bus has to wait until the first data transfer is completed. If the previous data transfer was already finished the execution proceeds without interruption. In this way, the FLIP has an asynchronous sequence control achieved by mutual sychronization of the individual processors with their joint data flow.

2.5.   Supporting Software

The FLIP is programmed in its machine languages called FAL (FIP Assembly Language) and PAL (PEP Assembly Language), respectively. The object modules are linked by FLINK (FLIP LINKage editor) which arranges the relative bus addresses within the objects to establish the desired logical structure on FLIP. Program execution is initiated and supervised by the FLEX (FLIP Executive) utility running on the host computer. FLEX controls all steps: The transfer of image data from a mass storage to FLIP, the FLIP execution, and finally the storing of the results produced by FLIP. Additional features of the control program FLEX are the capability to debug FLIP programs and to locate program errors by the use of trace routines, and to evaluate program efficiency by monitoring the activities of FLIP (FIP and PEP) by time measurements.

3.   FLIP IMAGE PROCESSING APPLICATIONS

Considering the special capabilities of FLIP described in the previous sections, one can easily realize that FLIP will find advantageous applications in the large field of picture processing. A great number of frequently used tasks like e.g. simple array operations for image filtering (linear or nonlinear), contour sharpening, and image enhancement imply parallel processing structures to be directly implemented with FLIP. Obviously, other time consuming image processing functions, e.g. local image correlation, image convolution or some gray level statistics are also suitable to be implemented on the FLIP, which already was successfully done. In the following a selection of processing structures will be given in order to demonstrate FLIP image processing, to give a performance evaluation, and to give a programmer's look to FLIP by means of programming techniques for FIP and PEP for two simple examples.

3.1.   Element Difference

The element difference method is a local operation used for generating of the first derivate of a picture F. It is a special case of the stroke difference method already described in the second section. The value of the difference $G(x,y)$ defined in both algorithms may be treated as a measure to characterize the gradient of the intensity distribution in this position (x,y). With the notation given in Figure 1 the element difference is given by:

$$Gx(x,y) = |F(2,1) - F(2,3)|$$

$$Gy(x,y) = |F(1,2) - F(3,2)|$$

$$G(x,y) = (Gx(x,y) + Gy(x,y))/2.$$

Considering the principles leading to the processing structure for the stroke difference, there is no difference in establishing a computing graph for the element difference function (Figure 9). This processing structure for the element difference is characterized by three pipelines with three processing stages each. Obviously, there are four processors needed at relatively equal load to perform the element difference for one submatrix position. However, this would be a bad system usage of FLIP as the other 12 IPs of FIP would remain unused. Due to the structural flexibility of FLIP, in our example it is possible to speed up computation by means of additional simultaneous processing, e.g. multiple organization of the simple processing structure for the element difference as far as processors are available. Figure 9 shows 12 IPs organized in three pipelines producing the element differences of three different submatrix positions at a time. The processor in the fourth stage of the processing cascade collects the results and outputs them for the purpose of storage or further processing, respectively. The data inputs for processors IP1-IP6 are fed from the PEP. Since PEP provides input data to FIP at a rate of up to 45 Mbyte, the loading of IP1-IP6 with input data can be considered as to happen in parallel.

Now, to show how simple it is to establish such a processing structure and to organize the data transports, the individual programs for the PEP-processor EP1 and the FIP-processors IP1, IP2, and IP3 are listed below:

```
        PEP - processor EP1:

START:
    MOV     F(2,1),IP1(A1)   ; Transfer pixel F(2,1) from buffer memory via bus A1 to
                             ; processor IP1
    MOV     F(2,3),IP1(B1)   ; same with F(2,3) and bus B1
    MOV     F(1,2),IP2(A1)   ; same with F(1,2) and bus A1 for processor IP2
    MOV     F(3,2),IP2(B1)   ; etc.
    IBR     P,START          ; Increment pointer to next window position and branch
                             ; to START
```

```
        FIP - processor IP1 and IP2:

START:
   SUB    A1,A2,SC1        ; Subtract gray levels of opposite pixels
   MOVM   SC1,OB           ; Output the absolute value
   BR     START            ; Goto START (next window-position)

        Processor IP3:

START:
   CLR    SC1              ; Provide ZERO-value
LOOP:
   ADD    IP1,IP2,OB       ; Add differences and output less significant byte
   ADC    SC1,OB           ; Output Carry (multiple precision add)
   BR     LOOP             ; etc.
```

Due to similarities between the programs listed above and the programs devoted for processors EP2, EP3, and IP5 - IP12, the latters are omitted here.

Finally, the output data rate is dictated by the processors with the heaviest load within the pipeline. The heaviest load found in the structure is 3 instructions per window position. Table 1 at the end of this section lists the selected FLIP image processing functions with their execution time.

3.2.   Two-Dimensional Convolution

Two-dimensional convolution is frequently used in spatial filtering, template matching, etc. It operates like a window operation and is described by the following expression:

$$G(x,y) = \sum_i \sum_j W(i,j) \; F\{i,j,(x,y)\}$$

where $W(i,j)$ is the weight matrix and $F\, i,j,(x,y)$ are the image points covered by the weight matrix at locations (x,y). Usually applied submatrix dimensions are ranging from 3 x 3 to 11 x 11 image elements. With FLIP it is possible to achieve one processing structure that is applicable for all submatrix dimensions considered. Nevertheless, processing structures will be more efficient if they are only devoted for one or perhaps two different submatrix sizes. Figure 10 shows a FLIP processing structure working on submatrix sizes of 5 x 5 and 7 x 7 pixels.

3.3.   Execution Times

Table 1 shows the execution times for the above mentioned image processing tasks. Not included in these times are the initialization and program load times.

Window Size	Element Difference	Stroke Difference	Two - Dim. Convolution
3 x 3	0.4 s	1.6 s	1.25 s
5 x 5	0.4 s	2.4 s	3.2 s
7 x 7	0.4 s	3.2 s	6.4 s

Table 1:   Execution times for 1024 x 1024 Pixels

4.   Conclusion

A powerful and suitable image processing system has been proposed and realized. The developed system possesses a flexible multiprocessor system for central processing and a high speed data exchange processor for fast and convenient data input/output. The FLIP is able to perform up to 64 MIPS (Million Instructions per Second) and can process image line lengths of up to 10 k Byte (depending on the submatrix size) and an unlimited number of lines.

Homogeneous operations in the field of image preprocessing, local image correlation and operations for feature extraction have been programmed. It was found, that in the most cases it takes less than 5 seconds to perform the considered image processing tasks on images with 1024 x 1024 pixels. The time required for processing is often less than the time that is needed to initialize the processing system and to take the image data to and from the storage device.

Due to its flexibility and processing power we believe FLIP will find practical application in other tasks of the same type, namely straightforward operations applied on a large set of data.

REFERENCES

CHEN, T.C.,   1971, "Parallelism, Pipelining and Computer Efficiency",
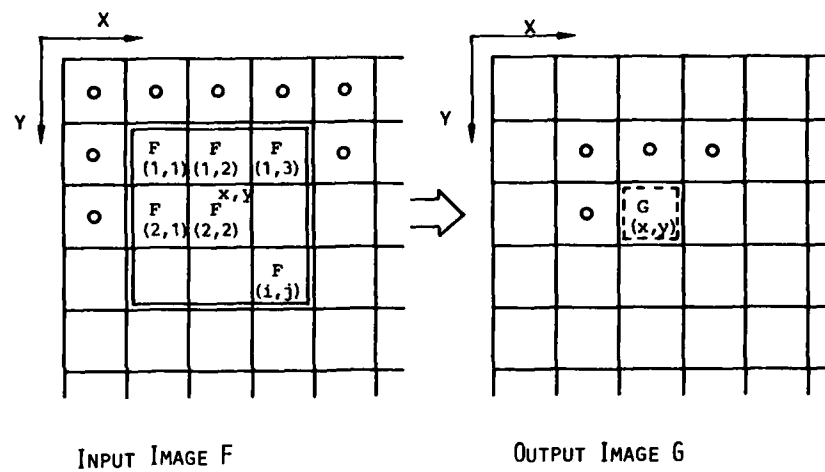   Computer Design, pp 69-74
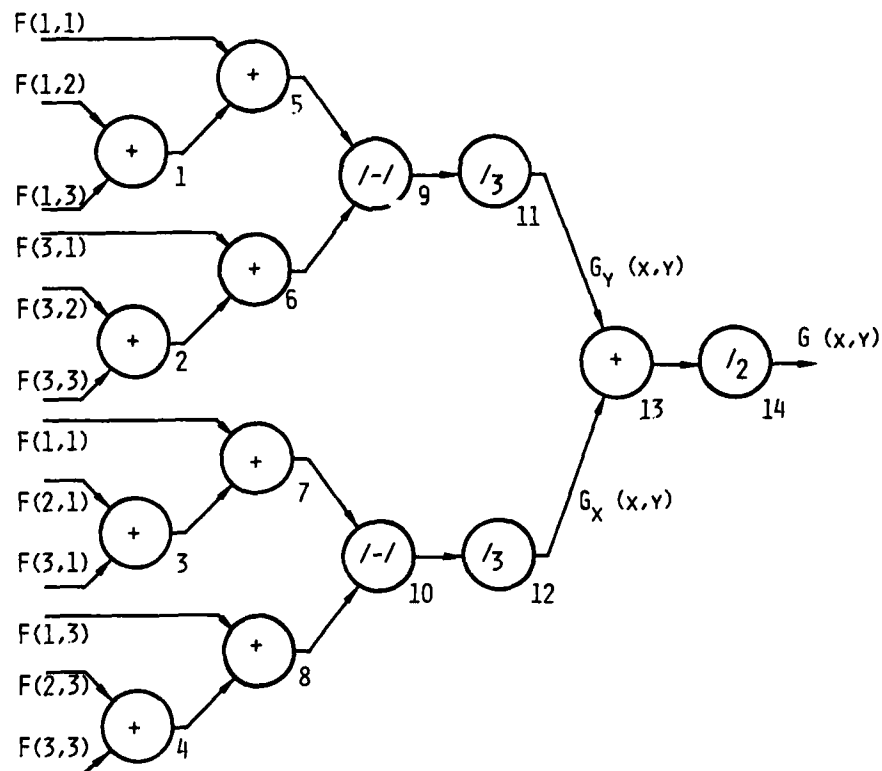
Figure 1:   Submatrix Notation for Window Operations

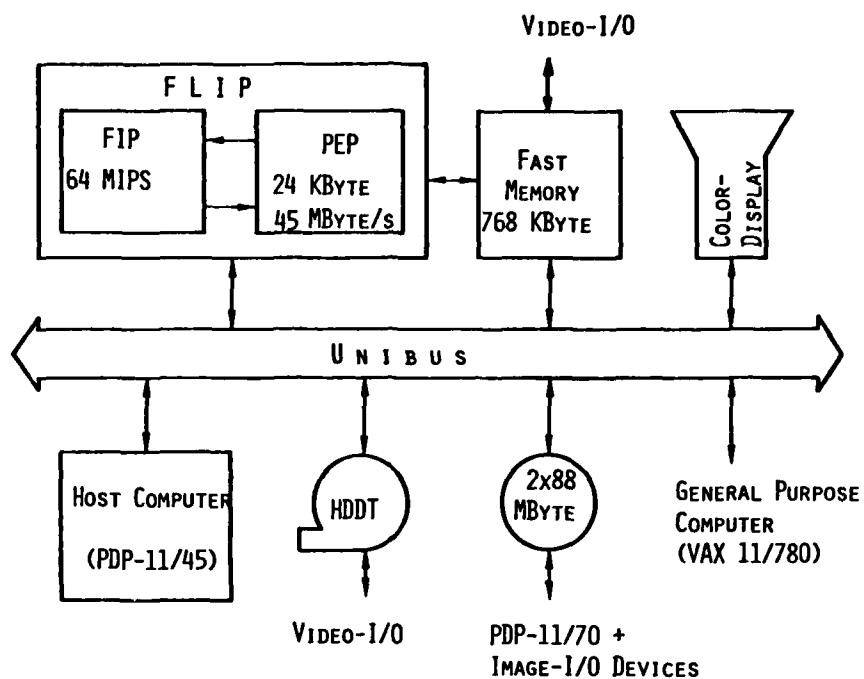Figure 2:   Computing Graph for Stroke Difference

VIDEO-I/O

FLIP

| FIP 64 MIPS | PEP 24 KBYTE 45 MBYTE/s | FAST MEMORY 768 KBYTE | COLOR-DISPLAY |

UNIBUS

| HOST COMPUTER (PDP-11/45) | HDDT | 2x88 MBYTE | GENERAL PURPOSE COMPUTER (VAX 11/780) |

VIDEO-I/O   PDP-11/70 + IMAGE-I/O DEVICES

Figure 3:   Image Processing System Configuration

| SINGLE DATA INPUT STREAM | MULTIPLE DATA STREAM | SINGLE DATA OUTPUT STREAM |

DISK — PEP ⇒ FIP — DISK

| IMAGE DATA STORAGE | IMAGE DATA MANAGEMENT | IMAGE DATA PROCESSING | IMAGE DATA STORAGE |

SDI   MDS   SDO

LINE BUFFERING,   PROCESSING
SUBMATRIX

INPUT IMAGE   OUTPUT IMAGE
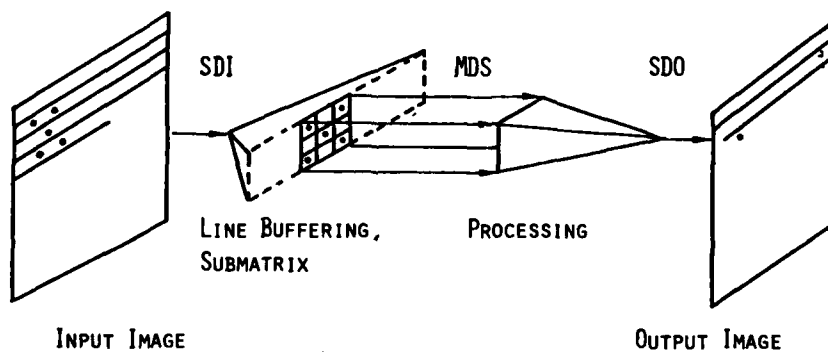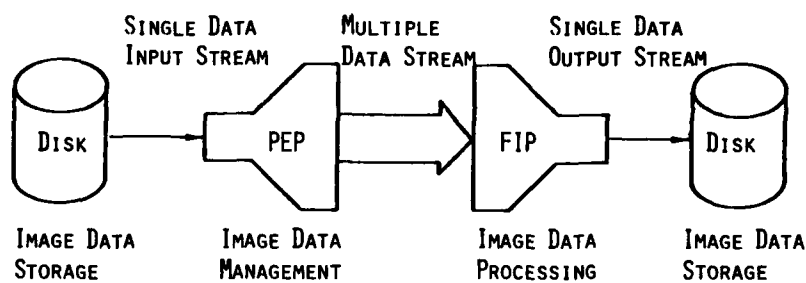
Figure 4:   FLIP Image Processing Data Flow
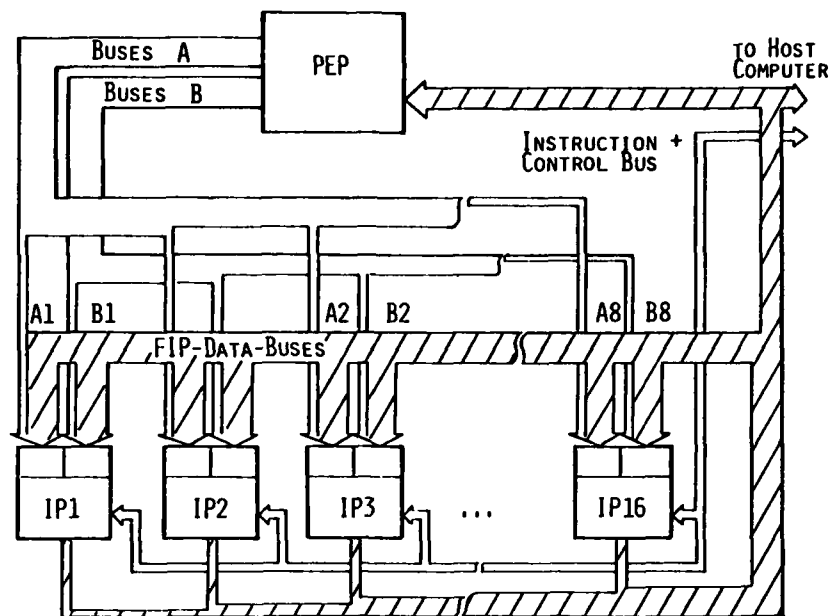
Figure 5: FLIP Bus System
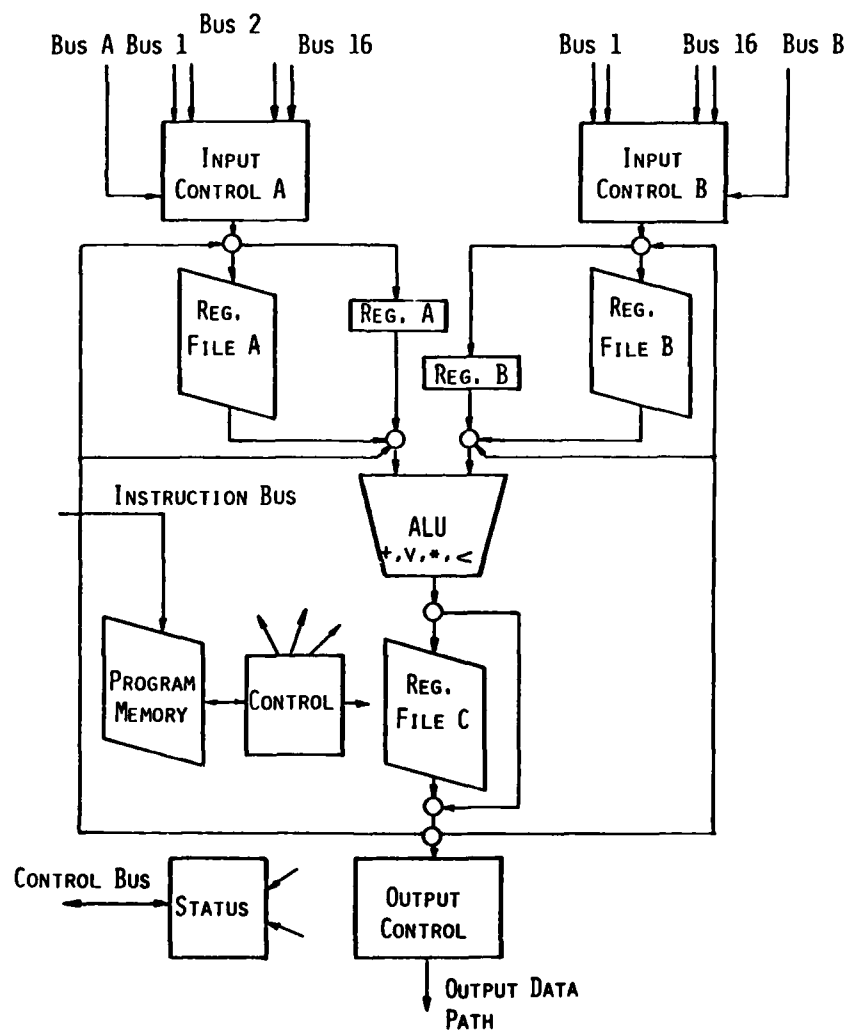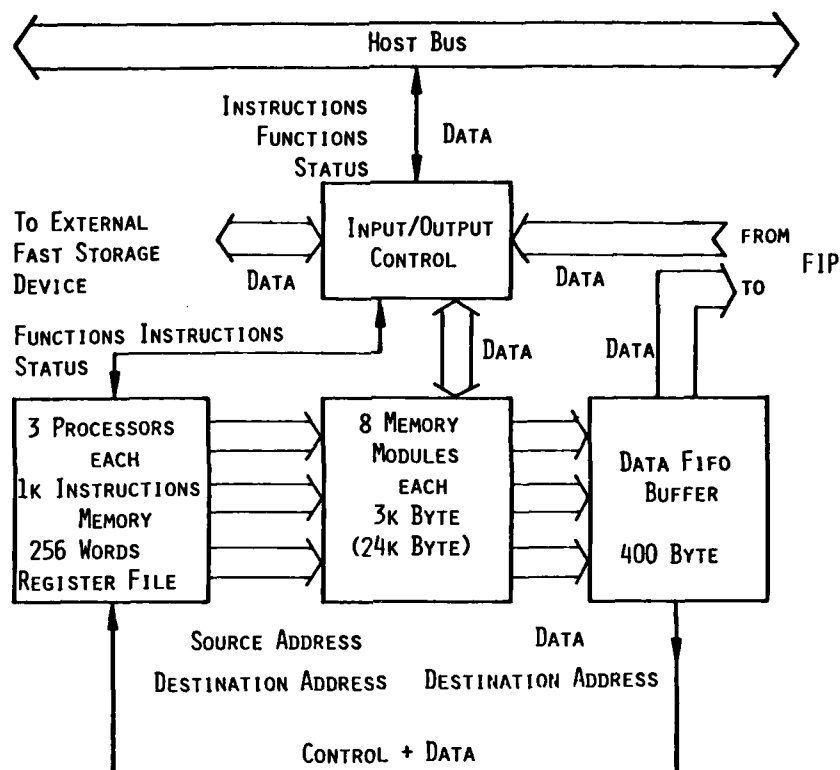
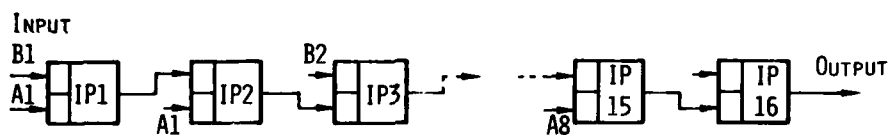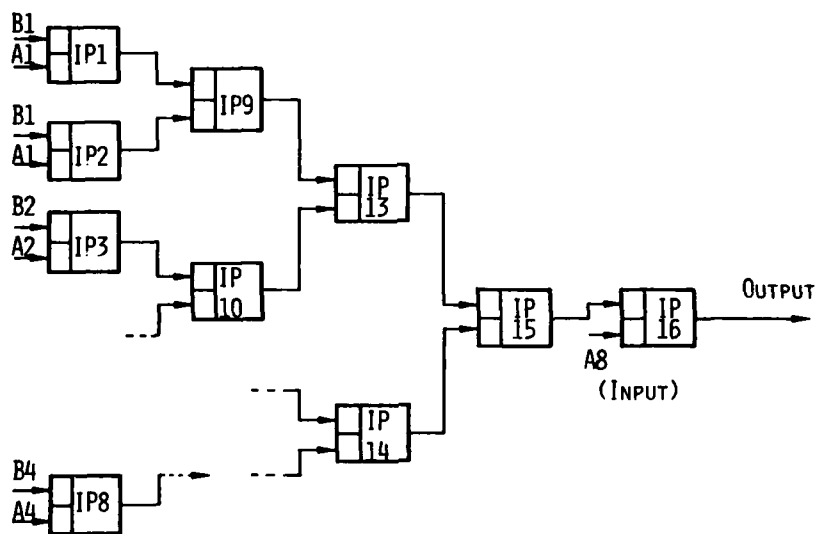

Figure 6: Structure of an Individual Processor

Host Bus

INSTRUCTIONS
FUNCTIONS
STATUS

DATA

INPUT/OUTPUT
CONTROL

TO EXTERNAL
FAST STORAGE
DEVICE

DATA

DATA

FROM

FIP

TO

FUNCTIONS INSTRUCTIONS
STATUS

DATA

DATA

3 PROCESSORS
EACH
1K INSTRUCTIONS
MEMORY
256 WORDS
REGISTER FILE

8 MEMORY
MODULES
EACH
3K BYTE
(24K BYTE)

DATA FIFO
BUFFER

400 BYTE

SOURCE ADDRESS

DATA

DESTINATION ADDRESS

DESTINATION ADDRESS

CONTROL + DATA

Figure 7:   Data Exchange Processor PEP

INPUT
B1
A1

IP1

IP2

A1

B2

IP3

IP
15

A8

IP
16

OUTPUT

PIPELINE USING 16 PROCESSORS

B1
A1

IP1

B1
A1

IP2

IP9

B2
A2

IP3

IP
10

IP
13

IP
15

OUTPUT

IP
16

A8
(INPUT)

IP
14

B4
A4

IP8

CASCADE USING 16 PROCESSORS

Figure 8:   Processing Configurations with FIP

15-10
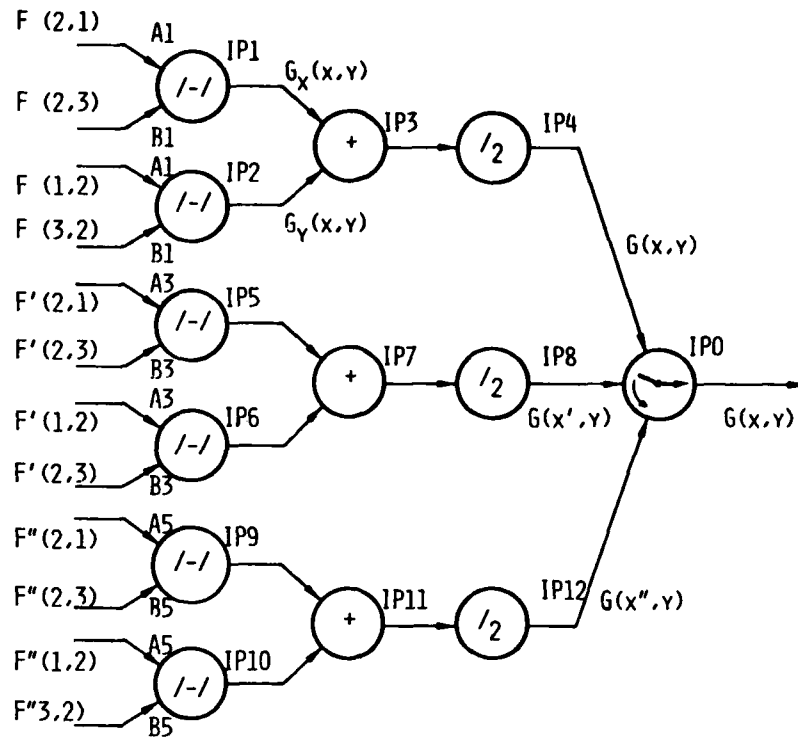
E.G.: $F'(1,2) = F\{1,2,(x',y)\}$; $x' = x+1$

$F''(1,2) = F\{1,2,(x'',y)\}$; $x'' = x'+1$

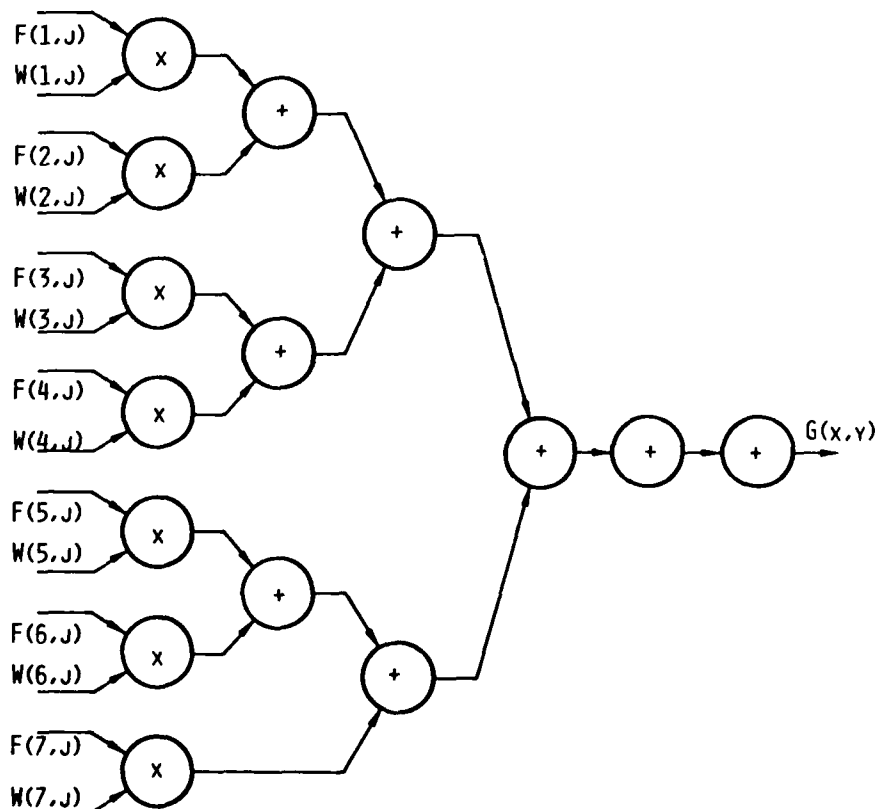Figure 9: Computing Graph for Element Difference

Figure 10: Computing Graph for Two Dimensional Convolution

AIDE A L'INTERPRETATION EN IMAGERIE SATELLITE MULTISPECTRALE

M. REBUFFET   Laboratoire de Traitement d'Images
Etablissement Technique Central de l'Armement
94114 ARCUEIL CEDEX
FRANCE

Cette contribution aborde plusieurs aspects de l'aide que les méthodes de traitement d'images sur ordinateur peuvent apporter à l'extraction de l'information contenue dans les images prises par un satellite d'observation.

Cette extraction a pour but d'établir une cartographie et une identification de planimétrie. Elle fait appel à des photointerprètes, pour lesquels on doit limiter le nombre de documents à consulter, et faciliter l'interface Homme-Machine.

Trois aspects sont développés dans cette contribution :

I   Amélioration d'une image multispectrale et réduction de la redondance,
II Correction de la géométrie d'une image par rapport à une référence,
III Extraction interactive des informations de type graphique.

Ces méthodes ont été mises en oeuvre au Laboratoire de Traitement d'Images de l'ETCA, sur une configuration de matériel de présentation et de traitement à haut degré d'interactivité.

I - AMELIORATION D'UNE IMAGERIE MULTISPECTRALE ET REDUCTION DE LA REDONDANCE

Le satellite LANDSAT produit des enregistrements sur bande magnétique de l'image multispectrale de la surface terrestre. Ces données font depuis longtemps déjà l'objet de recherches destinées à mieux connaître les ressources terrestres.

Dans les études qui font l'objet du présent rapport, nous avons utilisé ces données pour évaluer des méthodes de traitement de données multispectrales aériennes, et d'extraction ou détection d'objets d'intérêt militaire.

La finesse des images données par ce satellite est adaptée à des objets ou sites de grandes dimensions, supérieures à 50 m. Pour les objets plus petits, une extrapolation est nécessaire, mais les idées générales qui se dégagent de cette étude sont adaptables facilement et pourront aider à la conception de systèmes de traitement d'images de satellites d'observation militaire.

Le problème principal de la détection d'objets d'intérêt militaire est que l'information principale qu'on cherche à extraire est la perturbation que l'homme apporte au paysage naturel. Cette perturbation est faible en terme de pourcentage de surface de l'imagerie aérienne. Les traitements statistiques seront donc dans cette mesure délicats à utiliser.

La présente étude porte sur une méthode de traitement statistique, la transformation de Karhumen-Loeve dans l'espace chromatique, ou analyse factorielle qui permet d'obtenir théoriquement à partir de données multispectrales, des images décorrélées (dont les informations données sur chaque image sont de nature différente) et de meilleur contraste.

Cette transformation est très utilisée en Ressources Terrestres, mais où en général les données à détecter sont importantes en pourcentage surfacique.

La transformation de Karhumen-Loeve permet d'avoir de nouvelles images dont le rapport signal sur bruit est rangé suivant des valeurs décroissantes. Les photointerprètes en ressources terrestres voient en général apparaître des renseignements intéressants sur la deuxième image transformée, où la nature du sol apparaît et où les ombres ont presque disparues. La première image étant l'image "panchromatique" du paysage qui au contraire fait apparaître plutôt les ombres. Les militaires par contre sont intéressés a priori par les deux informations.

En conséquence nous avons étudié sur une zone où existait un certain nombre d'installations ou sites d'intérêt militaire, elles que zones industrielles, zones urbaines, voies ferrées, routes, aérodromes. Nous avons appliqué la transformation sur l'ensemble de l'image et sur un certain nombre de zones intéressantes. Nous avons mesuré la stabilité de cette transformation sur l'ensemble de l'image et obtenu des restitutions photographiques, de chaque image transformée.

ANALYSE FACTORIELLE

La méthode de l'analyse factorielle est une approche systématique mais linéaire d'un problème exprimé par les questions suivantes :

- Combien y a-t'il de variables indépendantes ?

- Comment les obtenir à partir des variables mesurées ?

Dans le cas de données naturelles qui possèdent un caractère statistique, on peut aussi formuler la question suivante :

- Quelle fraction de la variance des données néglige-t-on, lorsque, après avoir décorrélé les données, on ne conserve que m variables m < k ?

Admettons que nous ayons fait N expériences, dans lesquelles pour chacune on a enregistré k signaux expérimentaux. Les k sources de signaux, éventuellement entachée d'erreurs gaussiennes en première approximation, sont-elles indépendantes, corrélées ou redondantes ?

Soient $(x_i j)$ i = 1,N, J = 1,k les mesures ; soit $\mu_j = \frac{1}{N} \sum_{i=1,N} x_i j$ les moyennes de chaque signal et

$\sigma_j{}^2 = \frac{1}{N-1} \sum_{i=1,N} (x_i j - \mu_j)^2$ leurs écarts-types. On construit la matrice de corrélation entre les k si-

gnaux par :

$$c_{\ell,m} = \frac{1}{\sigma_\ell \sigma_m} \sum_{i=1,N} (x_{i,\ell} - \mu_\ell)(x_{i,m} - \mu_m).$$

Si les k signaux sont indépendants, cette matrice est diagonale. S'ils sont corrélés mais non redondants, cette matrice n'est pas diagonale, mais n'est pas singulière. Si les signaux sont redondants, la matrice est singulière, et se réduit à une matrice de rang inférieur. Voir figure 3 la représentation de ces trois cas lorsque k = 2.

Dans le cas de signaux gaussiens, a , les axes de l'ellipse parallèles aux axes du repère caractérisent l'indépendance. Dans le cas de signaux corrélés b , les axes sont obliques. Dans le cas de la redondance c l'ellipse se réduit à une droite. (voir figures 1a - 1b - 1c).

Pour obtenir les inclinaisons des axes principaux de cette ellipse, ou de l'hyperellipsoïde dans le cas général il suffit de diagonaliser la matrice de covariance et les vecteurs propres représentent l'importance relative des diamètres de l'ellipsoïde relatifs à chaque axe. Chaque axe représente un "facteur" ou "composante principale", chaque valeur propre mesure l'importance de ce facteur.

Par ailleurs si chaque mesure est entachée d'une erreur (cf 2.3.), et si ces erreurs sont indépendantes, ce qui correspond à la notion de bruit blanc, le rapport S/B que représente le rapport entre un facteur et la projection du bruit sur son axe, est proportionnel à la valeur propre correspondante.

Application aux données multispectrales

On applique la méthode précédente, les N expériences étant les N points (ou pixels) de l'image et les k signaux les k longueurs d'onde d'analyse (ici k = 4).

Donc faire l'analyse factorielle d'une image multispectrale c'est mettre en évidence des ensembles de quatre matrices de nombres représentant de nouvelles images dans lesquelles une "couleur" particulière est renforcée. La transformation de Karhumen-Loeve consiste à créer ces nouvelles images, où le rapport signal/bruit est optimisé pour la première image, puis pour la deuxième, celle-ci étant décorrélée avec la première, puis pour la troisième, décorrélée avec les deux premières etc...

Organigramme de l'analyse factorielle

Le programme doit exécuter plusieurs tâches :

a) Calcul des moyennes des données

b) Calcul des variances des données

c) Calcul de la matrice de corrélation

d) Recherche et tri des valeurs propres et vecteurs propres

e) Calcul de la matrice d'hyperrotation permettant de passer des variables mesurées aux nouvelles variables orthogonales.

APPLICATION A UNE IMAGE COMPLETE

Résultats et remarques

La corrélation entre les canaux 1 et 2 est la plus forte : 0,90 ; elle est en revanche très faible entre les canaux 1 et 4 : 0,21. La décroissance rapide des valeurs propres montre que le nuage statistique est très plat, et se développe essentiellement dans le plan formé par les deux premiers vecteurs propres.

L'épaisseur du nuage statistique autour de ce plan est très faible ce qui signifie qu'en ne conservant que les deux premières images transformées, on néglige seulement 4% de l'information statistique. La visualisation des deux dernières images transformées confirme le faible intérêt de ces deux dernières images, qui contiennent surtout du bruit.

Toutefois, le long du troisième axe, on reconnait en dehors du bruit une information qui représente le gauchissement du nuage statistique autour du plan défini par les deux premiers axes, et qui met en évidence certains détails à haute fréquence spatiale dans l'image.

Les deux premières images transformées

La première image transformée est obtenue par une pondération pratiquement identique des quatre canaux bruts. Elle a donc l'aspect de l'image panchromatique qu'on aurait obtenu avec un capteur à large bande spectrale (0,5 à 1,1μ). En effet l'écart angulaire dans l'hyperespace entre la combinaison linéaire calculée et la simple moyenne est égal à 0,1 rd.

Le rapport signal/bruit de l'image est deux fois plus grand que dans les images brutes, ce qui se concrétise par une nette atténuation des effets de trame existant dans les images LANDSAT.

Les points les plus clairs représentent les carrières, les digues, les lits asséchés des rivières, les zones à végétation éparse sur sol caillouteux ainsi que les structures en béton en particulier l'aérodrome d'Orange.

Les points les plus sombres représentent l'eau des rivières, les lacs, ainsi que la végétation arborescente.

La seconde image transformée est obtenue comme différence entre les canaux (0,7 - 0,8μ) et (0,8 - 1,1μ) et les canaux (0,5 - 0,6μ) et (0,6 - 0,7μ). On y voit donc apparaître en très clair l'ensemble des végétaux (forte réflectance dans le proche I.R. et faible réflectance dans le visible). Par opposition, et ceci est d'un grand intérêt pour le but visé, tout ce qui est minéral se distingue en sombre : eau des rivières, cailloux des lits des rivières, tuiles et métal des toits des constructions dans les villes, dalles de béton des usines, béton et bitume des pistes des aérodromes, voies de communications importantes, gares de triage. On remarquera en particulier que dans cette seconde image, on peut mesurer sans erreur la longueur des pistes des aérodromes, alors qu'aucune autre image (brute ou première transformée) ne le permet lorsque ces pistes n'utilisent pas le mˆeme matériau d'extrémité en extrémité.

L'application de la transformation de Karhumen-Loeve aux images multispectrales entières, ou fractions d'images de plus en plus petites a permis de mettre en évidence plusieurs résultats :

1. La perte d'information a priori (et non utile) que représente le passage de quatre images brutes à deux images tranformées est inférieure à 5%.

2. Ce qui est plus important c'est que l'information utile contenue dans les deux dernières images transformées est sémantiquement très faible. En effet, on voit se concentrer dans ces deux dernières images tout le bruit, et la notion de forme en est pratiquement absente.

3. La qualité des deux premières images transformées est nettement supérieure à celle des images brutes. Ces images sont débruitées. Elles rendent donc le travail du photointerprète plus sûr.

4. On peut attacher à chacune de ces deux images, une signification : image panchromatique pour la première et image séparant le minéral du végétal pour la seconde.

5. Lorsqu'on s'attache à des zones de petite taille, le plan qui définit les deux premières images transformées est d'orientation pratiquement constante. Ceci est très important, car toute rotation dans ce plan conserve la même information statistique, on pourra donc rechercher ensuite dans ce plan, la rotation qui au regard de certains critères apportera la plus grande information utile (au sens sémantique).

6. La stabilité de ce plan autorise à échantillonner dans une grande proportion le signal lorsqu'il s'agit d'en calculer les paramètres.

7. La notion de signature spectrale réduite à ce plan a un sens, ce qui ouvre la porte à une aide informatisée rapide et peu coûteuse pour les photointerprètes. En particulier, une classification n'utilisant que ces deux premières composantes est remarquablement significatives.

II - CORRECTION DE LA GEOMETRIE D'UNE IMAGE PAR RAPPORT A UNE REFERENCE

Un satellite d'observation donne de la fraction du sol qu'il observe une image déformée par son attitude et ses variations d'attitude, l'atmosphère et le relief. Lorsque ces déformations ne sont pas trop importantes, elles ne gènent en rien la photo-interprétation de la scène. Mais les déformations mêmes faibles empèchent toute comparaison automatique de deux vues prises à des dates différentes, ainsi que tout positionnement des éléments extraits sur un référentiel sol. Cette référence peut être soit une carte si on en dispose, soit une autre vue de la scène, choisie pour sa bonne qualité.

Si certaines causes de déformations sont modélisables et mesurables, elles seront préalablement corrigées de façon déterministe.

Pour les autres causes, seul un calage interactif de la scène sur la référence peut être utilisé.

C'est cette procédure qui est décrite ici, avec ses étapes successives et les méthodes utilisées.

ANALYSE ALGORITHMIQUE DE LA CHAINE DE RECALAGE

SELECTION DE LA ZONE

La chaîne interactive de rectification d'images accède à des données cartographiques ou radiométriques qui sont stockées sur des mémoires de rafraichissement et son limitées au format 512 x 512.

Une mise en coïncidence grossière des 2 fenêtres homologues est effectuée en sous-échantillonnant au pas de deux et en présentant à l'opérateur une fenêtre centrale de 256 x 256 de la référence ou l'image brute entière. En visualisant en alternance les 2 mémoires et en déplaçant un cadre 256 x 256 dans le plan graphique associé à l'image brute, l'opérateur extrait une fenêtre dont le centre coïncide à peu près avec le centre de la fenêtre de référence.

Les coordonnées du coin supérieur gauche du cadre graphique sont stockées sur fichier et lues ensuite par le programme de transfert de bande à disque. Les données de travail sont dès lors stockées sur les mémoires de rafraichissement de la console de visualisation.

CHOIX DES "AMERS" SUR LA REFERENCE

Par l'intermédiaire d'un curseur carré de 32 x 32, l'opérateur sélectionne le domaine inscrit dans le curseur en tant "qu'amer" de l'image de référence. Les statistiques associées à cet "amer", la distribution en radiométrie, ainsi que la localisation du coin supérieur gauche sont stockés pour "l'amer" courant sur le fichier AMERS.

Il faut souligner ici un problème associé à la désignation par curseur sur le plan image de la console de visualisation :

La localisation en ligne comporte une incertitude de 1 pixel.

Le nombre "d'amers" saisis conditionne la qualité de la régression des mesures sur la fonction polynomiale modèlisant les déformations.

POINTAGE DES "AMERS" SUR LA SCENE

On présente successivement à l'opérateur en 3 passes l'image de référence dans la moitié supérieure de l'écran et l'image brute dans la moitié inférieure, ainsi que les "amers" sélectionnés. Pour chaque "amer" visualisé comme un carré inscrit dans un disque, l'opérateur doit pointer la zone correspondante de l'image brute.

Un affinage de cette procédure de pointage peut être réalisé sur des données image en cherchant le maximum de la matrice de corrélation sur des fenêtres de 32 x 32. La corrélation est calculée globalement à l'aide des transformées de Fourier discrètes par la formule :

$$- \text{CCF} = \frac{1}{N} \mathcal{F}^{-1} \left(\mathcal{F}_{(f)} \times \mathcal{F}_{(g)} \right)$$

où N est un facteur de normalisation et $\mathcal{F}$ l'opérateur transformée de Fourier.

Le pic de cette matrice est recherché dans un voisinage 16 x 16 du centre et s'il existe, sa position est donnée par les coordonnées barycentriques dans une fenêtre 5 x 5 autour du pic.

La dynamique de la matrice de corrélation est un indicateur de qualité sur "l'amer" courant qui, s'il est inférieur à un certain seuil, peut conduire au rejet de "l'amer" par l'opérateur.

Le calcul des transformées discrètes sur une fenêtre 32 x 32 introduit des erreurs de phase lorsque des discontinuités aux frontières existent, l'opérateur peut donc optionnellement appliquer une fonction fenêtre (i, e, Gaussienne) aux zones traitées.

CALCUL DE LA FONCTION DE DEFORMATION

Les couples de points homologues sont ensuite utilisés pour faire un ajustement polynomial aux moindres carrés sur l'expression :

$$X(xy) = b(1)x+b(2)x^2+b(3)y+b(4)y^2+b(5)xy+b(6)xy^2+b(7)x^2y+b(8)x^3+b(9)y^3+b(10)$$

$$X(xy) = c(1)x+c(2)x^2+c(3)y+c(4)y^2+c(5)xy+c(6)xy^2+c(7)x^2y+c(8)x^3+c(9)y^3+c(10)$$

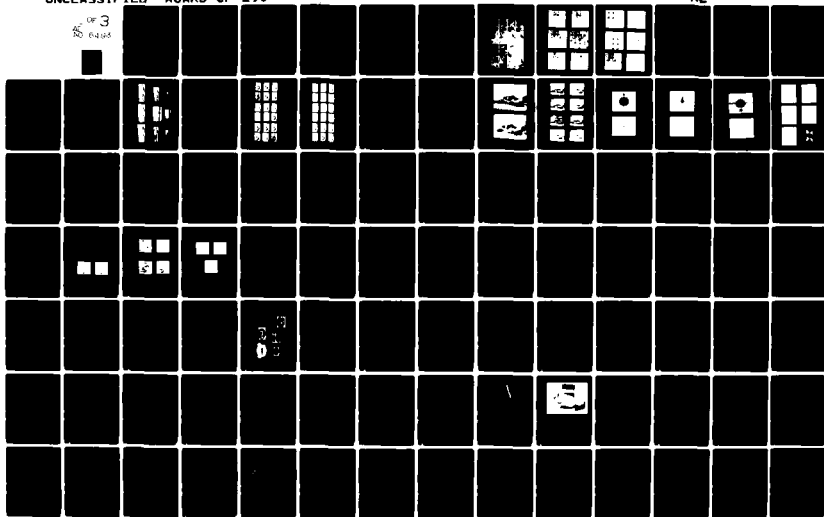ou sur des expressions comparables pour des polynômes de degré 1 ou 2.

Dans cette expression, X et Y sont les coordonnées écran du centre de la fenêtre image brute associée à "l'amer", x et y les coordonnées du centre de "l'amer", et b, c les coefficients de la transformation.

Pour des déformations localement importantes, une version améliorée de cette procédure de modélisation consisterait à calculer les transformées des "amers" de référence avant de faire la corrélation numérique avec les imagettes de l'image brute, de façon à obtenir une meilleur fiabilité dans la recherche du pic.

RECTIFICATION DE L'IMAGE

Lors de la rectification, on calcule les transformées des noeuds d'une grille régulière de l'image de référence et on associe au point transformé ou "point père" une radiométrie déduite par interpolation des radiométries de son voisinage dans l'image brute.

Si $\vec{P}$ désigne la transformation polynomiale dont les coefficients b et c ont été calculés dans l'étape de modélisation, la rectification comprend 3 phases :

- calcul du transformé $\vec{P}(i,j)$ pour chaque point (i,j) de la grille de référence :

- détermination du voisinage de l'image brute associé au point père et du jeu de coefficients d'interpolation correspondant à la position de ce point dans le sous-maillage d'une maille élémentaire,

- calcul de la radiométrie du point transformé par interpolation bilinéaire dans un voisinage 4 x 4 autour du point père.

Comme la rectification est très coûteuse en temps de calcul, la programmation de ces 3 phases a été optimisée :

- Pour chaque ligne d'un bloc d'image, les coordonnées des points transformés sont calculées itérativement à partir des coordonnées du premier point. La représentation interne de ces coordonnées est en format virgule flottante pour les transformations de degré > 2, et en entier étendu pour les transformations linéaires : cette dernière représentation permet d'éviter d'avoir recours aux opérateurs virgule flottante peu performants.

- Un sous-maillage de 8 x 8 pour une maille élémentaire offre une précision suffisante pour les interpolations bilinéaires ; on doit donc stocker 64 jeux de matrices de coefficients b de taille 4 x 4, et accéder pour le point courant d'une ligne au jeu correspondant.

  Ces coefficients sont stockés comme des entiers et comprennent donc un facteur de normalisation.

- La radiométrie du point transformé est calculée en sommant les contributions des points du voisinage, pondérées par le coefficient associé, puis en normalisant le résultat. Les calculs se font en représentation entier étendu, c'est-à-dire, sur 32 bits.

  Les données radiométriques de l'image brute accédées par cette étape d'interpolation, sont stockées dans un tampon circulaire mis à jour après le traitement de chaque ligne. Une table de pointeurs, préparée lors de l'initialisation, permet d'accéder au premier point d'une ligne donnée au sein du tampon circulaire.

Interpolation radiométrique

Les coefficients d'interpolation bilinéaire sont calculés en prenant comme hypothèse la conservation de l'énergie du signal lors de cette interpolation par une fonction cubique. On supposera en outre l'indépendance en x et en y de la fonction d'interpolation, la dérivation des coefficients étant présentée dans une dimension.

Si $f(x)$ est la fonction d'interpolation, $F(x)$ sa primitive, l'hypothèse de conservation d'énergie du signal sur les 4 intervalles conduit au système d'équations :

$$a_1 = F(-1) - F(-2)$$
$$a_2 = F(0) - F(-1)$$
$$a_3 = F(1) - F(0)$$
$$a_4 = F(2) - F(1)$$

avec

$$f(x) = \alpha x^3 + \beta x^2 + \gamma x + \delta$$
$$F(x) = \alpha \frac{x^4}{4} + \beta \frac{x^3}{3} + \gamma \frac{x^2}{2} + \delta x + \varepsilon$$

La résolution de ce système d'équations linéaires conduit au résultat suivant :

$$\alpha = \frac{1}{6}(-a_1 + 3a_2 - 3a_3 + a_4)$$

$$\beta = \frac{1}{4}(a_1 - a_2 - a_3 + a_4)$$

$$\gamma = \frac{1}{12}(a_1 - 15a_2 + 15a_3 - a_4)$$

$$\delta = \frac{1}{12}(-a_1 + 7a_2 + 7a_3 - a_4)$$

La fonction d'interpolation, issue de cette hypothèse de conservation d'énergie se déduit des coefficients associés aux radiométries des points du voisinage comme suit :

$$f(x) = a_1(-\frac{1}{6}x^3 + \frac{1}{4}x^2 + \frac{1}{12}x - \frac{1}{12}) + a_2(\frac{1}{2}x^3 - \frac{1}{4}x^2 - \frac{5}{4}x - \frac{7}{12}) + a_3(-\frac{1}{2}x^3 - \frac{1}{4}x^2 + \frac{5}{4}x + \frac{7}{12})$$

$$+ a_4(\frac{1}{6}x^3 + \frac{1}{4}x^2 - \frac{1}{12}x - \frac{1}{12})$$

Cette fonction comprend 4 composantes qui satisfont au critère de continuité aux bornes (i.e. P2(1/2) = P4(1/2)). Elle est calculée pour un maillage 8 x 8 autour de l'origine et stockée dans un fichier.

Extension de l'algorithme de rectification aux fortes déformations

Pour de faibles déformations, les données "image brute" sont stockées ligne par ligne dans le tampon circulaire : ce tampon est mis à jour après rectification d'une ligne, par libération des lignes non utilisées et remplissage par des lignes nouvelles d'indice plus élevé.

Dans cette méthode, la taille du tampon déterminera l'angle maximal de rotation locale d'une ligne. Sur le "Mitra 15", on est ainsi limité à 22 lignes d'image, correspondant à une rotation de moins de 3 degrés.

L'extension de l'algorithme à des déformations importantes consiste à effectuer le traitement par pavés au lieu de le faire séquentiellement sur une ligne de la mémoire. La taille des pavés est ici conditionnée par les options disponibles sur la "Trim" pour accéder à une zone mémoire en jouant sur le facteur de zoom et le numéro de partition. Le programme mis en oeuvre utilise des pavés de 256 colonnes qui peuvent être accédés en indiquant leur numéro de ligne au sein d'une partition.

Le traitement se déroule donc sur 2 bandes verticales de 256 colonnes successivement, les données "image brute" étant stockées dans le tampon circulaire et mis à jour dynamiquement pour économiser l'espace utilisé. Les transformées des frontières verticales de la bande courant servent à définir pour une ligne les valeurs de l'image brute à stocker dans le tampon, le pointeur de début de ligne étant chaque fois calculé et conservé dans la table des pointeurs ligne.

Cette gestion dynamique de la table des pointeurs offre l'avantage que le programme même de rectification reste inchangé, l'implantation complexe des données ne se répercutant que sur la gestion de ces pointeurs et sur le calcul des coordonnées du premier point d'une ligne avant la boucle itérative sur les colonnes.

En partitionnant l'image en 2 bandes verticales, on pénalise le temps de traitement d'une lecture supplémentaire de toute l'image. Les rotations locales peuvent ainsi atteindre des valeurs de l'ordre de 2,22/256 soit ≃ 12 degrés.

III - EXTRACTION INTERACTIVE DES INFORMATIONS DE TYPE GRAPHIQUE

On décrit ici une méthode de détection interactive des traits sur des images satellite de la surface terrestre. Ces traits représentent en général des routes, des autoroutes, des rivières, des canaux ou des voies ferrées, donc des voies de communication, leur forme et leur fonction étant d'ailleurs étroitement liées. C'est cette fonction de communication qui fait l'intérêt stratégique d'une telle détection.

L'observation d'une image satellite (ou d'une carte de même échelle) montre que la partie humaine (c'est-à-dire artificielle) de la surface représentée prend la forme d'un graphe dont les noeuds sont les villes et les branches les différentes voies de communication entre ces villes ; les grandes masses homogènes de l'image sont naturelles : surfaces d'eau, forêts, zones agricoles. Les villes sont relativement bien détectées par les méthodes multispectrales, par contre le suivi des routes est beaucoup plus difficile à établir par ces seules méthodes pour des raisons de finesse ou de contraste ; pour éviter un tracé en pointillé l'algorithme de détection a besoin de considérations géométriques et structurelles qui l'"aident" à retrouver le tracé cherché.

Des recherches sur la détection des traits ont déjà été menées dans plusieurs directions. Sur les tracés binaires (chambre à bulles). La théorie des graphes sert à la construction des lignes par croissance. Sur des images à plusieurs niveau de gris, plusieurs méthodes ont été utilisées : masquage sur une image de gradients, détermination des composantes connexes ayant un rapport surface sur périmètre très petit ou enfin détection de segments suivi d'algorithmes de jonction dans des directions provilégiées.

La méthodes présentée est un suivi de trait ou plus exactement le suivi d'un évènement à une dimension se propageant sur le domaine de l'image. Pour un trait cet évènement est schématisé par la fonction créneau (⎍), coupe du trait suivant une ligne ou une colonne de l'image. Ce suivi est initialisé par l'opérateur qui désigne le point de départ du trait et sa direction. Le suivi est alors effectué segment par segment.

A chaque étape, l'algorithme recherche ligne par ligne et par corrélation le point où le signal se rapproche le plus de l'évènement en conservant au mieux la direction précédente. Par régression on obtient deux nouvelles directions du trait à court et à long terme. L'écart entre ces directions modifie la vitesse de propagation du suivi et est une indication de la courbure du trait. Ce travail est effectué dans une sous zone de travail orientée dans une des 4 directions principales de l'image de manière à ce que le trait soit presque perpendiculaire aux lignes de la sous zone, lorsque le trait s'écarte trop de cette direction, une nouvelle orientation est prise.

Initialisation

L'opérateur désigne sur l'écran le point de départ du trait et un second point fixant la direction du trait et l'avancement initial de l'algorithme.

Sous-zone de travail

La sous-zone de travail est un carré de 64 x 64 pixels, orientée suivant l'une des 4 directions principales de l'image : verticale descendant ou ascendant, horizontale droite ou gauche. A chaque étape un test de débordement de zone est effectué, au cas de dépassement une nouvelle sous-zone est extraite de l'image.

Le travail de recherche du trait est effectué dans le repère lié à cette sous-zone, les changements de repère sont effectués au début et à la fin de la recherche, à chaque étape.

Cône de visée

Le cône de visée est le triangle défini par :

- le sommet : le dernier point du tracé M

- la hauteur issue de M : l'avancement ν

- la direction de la médiane issue de M : la direction θ du dernier segment tracé

- la longueur du côté opposé à M : $6\phi/(\nu + 1)$

Le cône de visée est le domaine de recherche du nouveau point du trait. L'angle d'ouverture du cône définit la variation maximale de direction du trait, il est décroissant en fonction de l'avancement ; en effet un faible avancement est lié à une courbure forte donc à une possibilité de forte variation de direction.

Modèle de corrélation

Un modèle du trait (pour le tracé passé) est calculé par moyennage des profils du trait pour les lignes du tracé précédent ayant une corrélation assez forte signal-modèle.

Le modèle est calculé en recalant chaque ligne de manière à faire coïncider les positions du centre du trait à chaque ligne. Le centre du trait est corrigé de manière à coïncider avec un extremum du signal.

Le modèle est la fonction C définir par :

$$C(x) = \left[\sum_{\text{ligne } k} s_k\ (p_k - \ell + x)\right]/n \qquad x \in [1, 2\ell]$$

où s_k est le signal à la ligne k

ℓ la 1/2 largeur du modèle est le centre du trait à la ligne k

$p_n r$ le nombre de lignes bien corrélées à l'étape précédente

Pour diminuer l'influence du bruit et augmenter l'inertie du modèle, le modèle de corrélation utlisé est finalement :

$$C = (C + Cp)/2 \text{ où Cp est le modèle de l'étape précédente.}$$

Corrélation

La corrélation est effectuée ligne par ligne. On recherche sur le segment de droite compris dans le cône de visée le point où le modèle du trait et le signal sont le plus proches (le modèle étant centré sur ce point). Soit p l'abscisse du point testé :

. Corrélation à gauche :

$$corg(p) = \sum_{x=1}^{\ell} \alpha(x)/s(p-\ell+x)-c(x)/$$

. Corrélation à droite :

$$cord(p) = \sum_{x=\ell+1}^{2\ell} \alpha(x)/s(p-\ell+x)-c(x)/$$

où ℓ est la 1/2 largeur du modèle

c le modèle

α une fonction poids qui valorise la contribution du centre du trait.

On a alors la corrélation au point p :

$$cor(p) = max \ (corg(p) + cord(p)) + 2 \ min \ (corg(p) + cord(p))/3$$

On cherche alors le point de la ligne ayant la meilleur corrélation soit le min de cor(p).

$Mac \ 1 = min \ ((cor(p)) \ p\epsilon|c_1c_2|)$ $Ipic \ 1 = p \ tg \ cor(p) = Mac \ 1$

$Mac \ 2 = min \ ((cor(p)) \ p\epsilon|c_1c_2|-Ipic1)$ $Ipic \ 2 = p \ tg \ cor(p) = Mac \ 2$

Le maximum de corrélation pour la ligne est alors :

Mac = Mac 1 pour le point Ipic = Ipic 1 sauf si Mac 2 est très proche de Mac 1 et Ipic 2 plus proche du pic de la ligne précédente que Ipic 1, alors Mac = Mac 2 et Ipic = Ipic 2.

Ceci est effectué pour les ν lignes du cônes de visée, on obtient donc un tableau :

$Ix(k) = Ipic$ pour la ligne k si Mac < Scor

$Ix(k) = \phi$ pour la ligne k si Mac $\geqslant$ Scor

où Scor est un seuil variable calculé d'après le pourcentage de bonne corrélation demandé par l'opérateur et fonction de la largeur du trait et du contraste.

On calcule enfin le taux de bonne corrélation :

$$tx = \frac{1}{\nu} \ (\text{\N} \ (ligne \ k \ tg \ Ix(k) > \phi))$$

Calcul de la direction du trait

On cherche les droites de régression des Ix(k) soient :

D_2 droite de régression des Ix(k) $k = 1, \nu/2$

D_1 droite de régression des Ix(k) $k = 1, \nu$

La droite cherchée est celle passant par M et dont la direction minimise la somme des carrés de distances à la droite des Ix(k).

On obtient 2 directions de visée : ϑ_1 à long terme et θ_2 à court terme, d'où le calcul de Id paramètre indiquant la stabilité directionnelle :

$$Id = \frac{s_1}{|tg\theta_1 - tg\theta_2|} \qquad où \ s_1 \ est \ un \ seuil \ donné$$

Id mesure la distance à partir de M telle que D_1 et D_2 ne soient pas écartées de plus de s_1.

Avancement

L'avancement ν va déterminer le nouveau cône de visée c'est-à-dire la possibilité d'accélération du tracé dans l'étape suivante :

Le nouvel avancement ν' est :

$$\nu' = Tx(Id + \nu/2)$$

Pour éviter les accélérations trop fortes on pondère cet avancement par un coefficient d'inertie Ks

$$Ks = 12/(\nu' + 1) \qquad et \quad \nu' = \nu' + Ks$$

Le segment tracé a alors pour origine le dernier point tracé et pour extrémité le point :

$(k,Ix(k))$ où k est la ligne tg

$k = \max_{k} \ ((k \ tg \ Ix(k) > \phi \ et \ k \leqslant \nu'/2))$

Changement de direction

Une trop grande valeur des paramètres θ_1 et θ_2 détermine un changement de direction, pour cela on teste la valeur :

$$T = |\theta_2| + 1-25|\ell_1|$$

$T < 2,5 \rightarrow$ pas de changement de direction

$T > 2,5 \rightarrow$ changement de direction

Soient M le point trouvé, MP le point précédent, MPP le point précédent MP et θp l'angle de direction précédent :

Si $\theta p < 1$ on pose M1 = MP, M2 = M et on reprend l'initialisation d'un trait avec changement de direction de la sous-zone de travail.

S $\theta p > 1$ on pose M1 = MPP, M2 = MP la réinitialisation a lieu avec un retour en arrière.

Arrêt du tracé

L'arrêt du tracé se fait dans les cas suivants :

(i)   – La direction θ_2 est extérieure au cône de visée.

(ii)   – L'avancement est trop faible (en général car $\theta_1 \neq \theta_2$).

(iii) – L'algorithme ne trouve pas de point valable dans le demi cône de visée.

(iv)   – Le trait sort des limites de l'image.

(v)   – Le nouveau modèle du trait a un contraste trop faible.

(vi)   – Le taux de bonne corrélation Tx est trop faible.

Dans les cas (iii) et (vi) il est prévu une nouvelle tentative de recherche du trait : si $|\theta_1|$ et $|\theta_2|$ sont assez grands on effectue un essai de tracé dans la direction voisine en adoptant la stratégie utilisée pour le changement de direction (avec retour en arrière).

A la fin du tracé on rend la main à l'opérateur qui définit la suit du programme.

Cette méthode permet le tracé de traits d'épaisseur quelconque et d'adapte aux changements de contraste au cours du suivi du trait, c'est son principal intérêt lié à sa rapidité due au caractère mono-dimensionnel du traitement. Le tracé du trait s'effectue à vitesse variable, rapide quand le tracé est linéaire, plus lent quand sa direction varie. Les problèmes sont liés à la gestion de la direction, soit dus au manque d'adaptation de la vitesse du tracé à la courbure du trait, soit dus à une direction du trait intermédiaire entre deux orientations de la sous-zone de travail (qui n'a que 4 directions possibles). Une amélioration serait donc apportée par la considération de 8 orientations pour la sous zone de travail (k x $\frac{\pi}{4}$ au lieu de k $\frac{\pi}{2}$) mais cela alourdirait considérablement le programme. Il est certain toutefois que l'ordinateur n'ayant pas la capacité de synthèse de l'oeil humain - qui "construit" le trait d'après une vision globale de l'image - Toute méthode de détection des traits se heurte au problème d'une détection purement locale, toujours partielle.
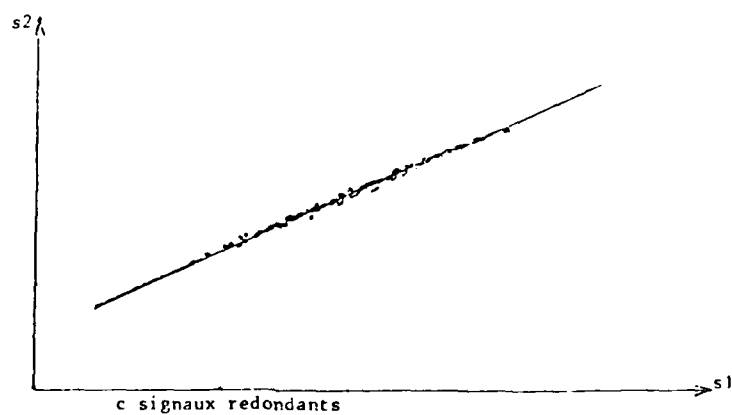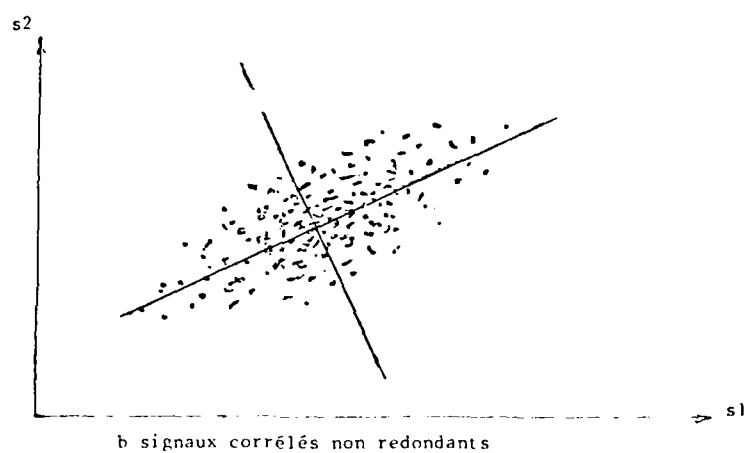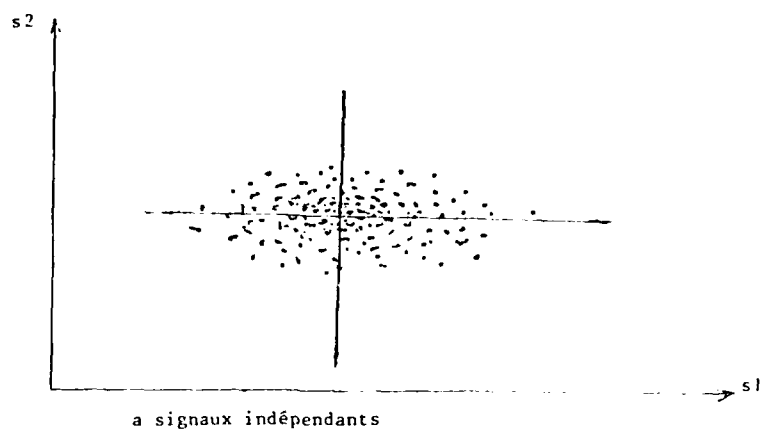
FIGURE   1

a signaux indépendants

b signaux corrélés non redondants

c signaux redondants

FIGURE 2

En haut les canaux bruts

En bas les canaux transformés

REGION D'AVIGNON (Sud-Est de la France)
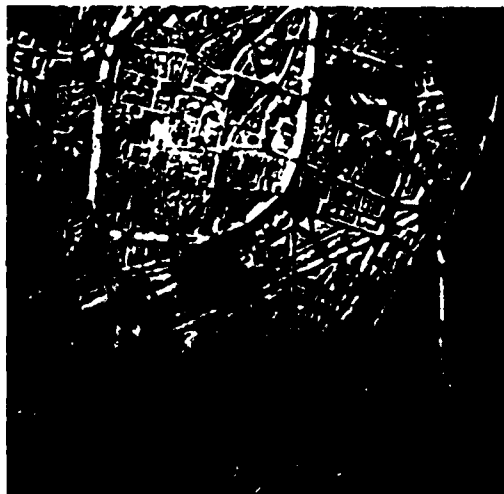
FIGURE 3 Correction géométrique

Photo 1   Image de Référence

Photo 2   Image déformée
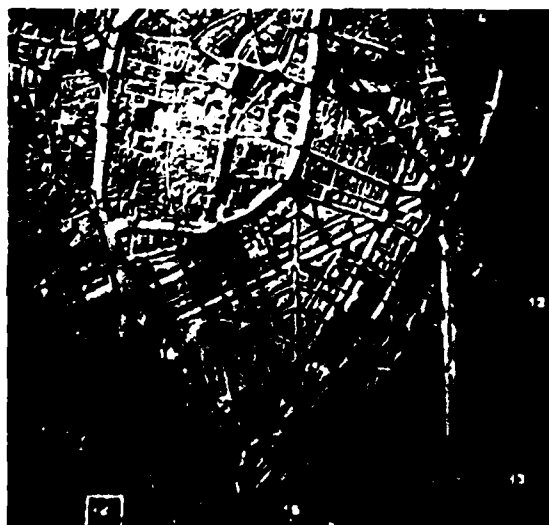

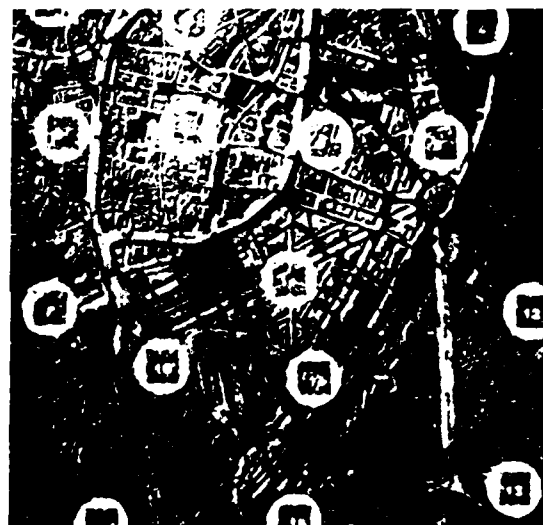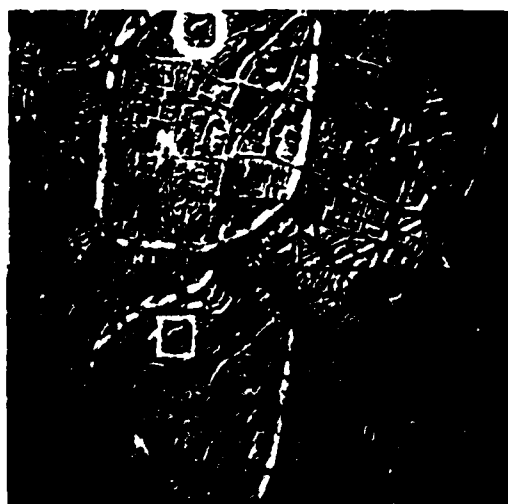Photo 3   Choix d'un Amer

Photo 4   Tous les Amers choisis

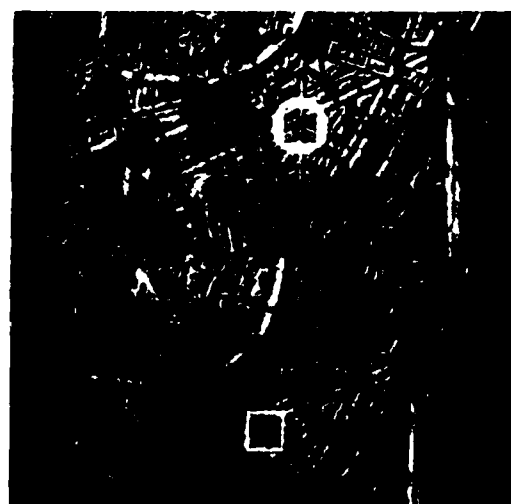Photo 5   Pointage d'un Amer

Photo 6   Pointage d'un Amer

FIGURE   6   Correction géométrique
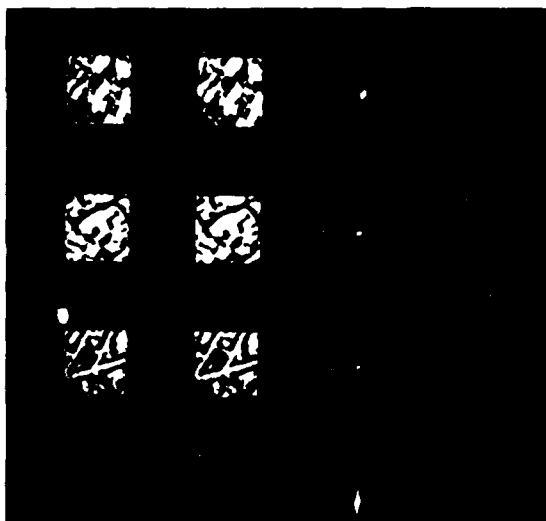
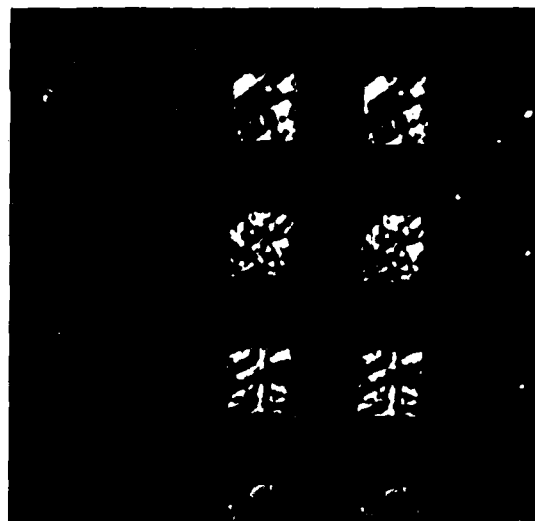Photo 7   Affichage Amers 1 à 4
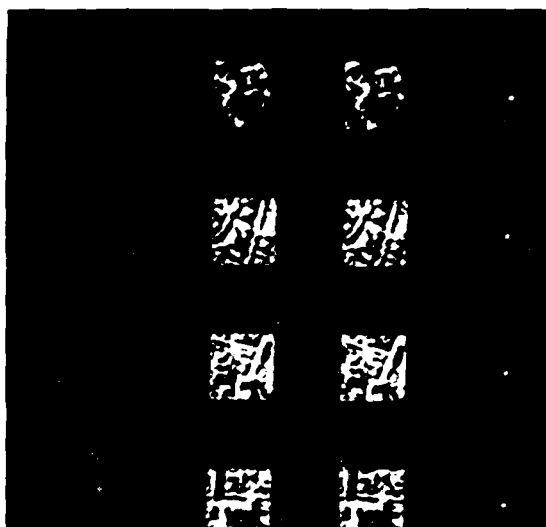
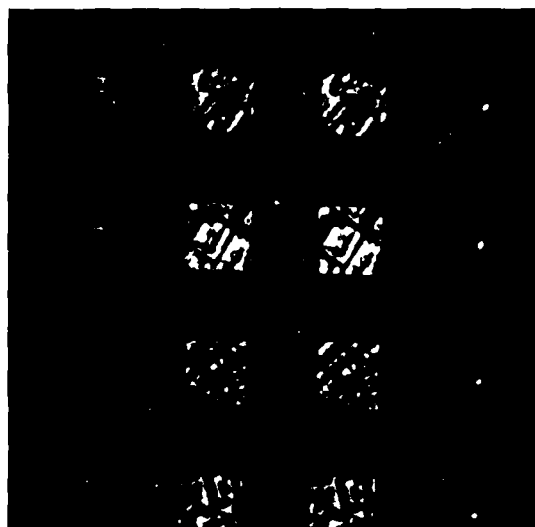Photo 8   Affichage Amers 9 à 12

Photo 9   Affichage Amers 5 à 8

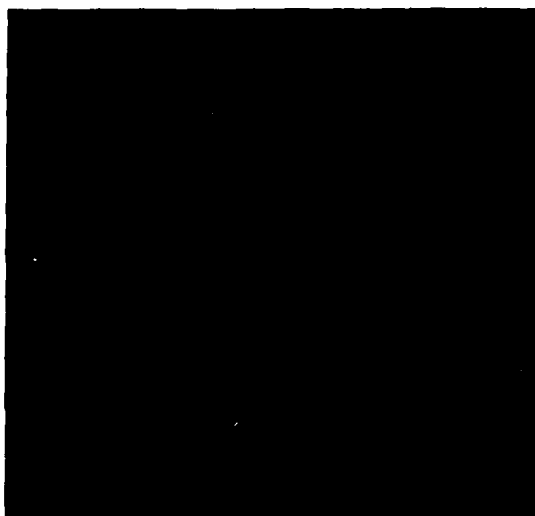Photo 10   Affichage Amers 13 à 16

Photo 11   Image corrigée

Photo 12   Erreur résiduelle

RESEARCH INTO METHODS OF IMAGE PROCESSING FOR TARGET ENHANCEMENT AND DETECTION

by

D.B.Duke, A.J.Fryer & P.A.Bird
British Aerospace Dynamics Group
Infra-Red Equipment Division
Hatfield, Hertfordshire, England.

1. INTRODUCTION

This presentation is aimed at showing some of the image processing techniques developed and used by the infra-red equipment division of British Aerospace at Hatfield. Work in image processing has been going on in the division since the early 1970's and has concentrated initially on techniques for optimising the performance of light weight thermal imagers for battlefield surveillance. Recent work however has been aimed at developing methods of enhancing targets with respect to their backgrounds for ease of detection and recognition. This has ultimately led on to our current work to target acquisition and tracking. The paper is divided into two sections:--

(i) Methods of Sensitivity Enhancement

(ii) Methods of Spatial Resolution Enhancement.

Use of these methods for improving image interpretability and automatic target detection are briefly discussed.

2. SENSITIVITY ENHANCEMENT

The Company has been interested mainly in thermal infra-red systems where the noise has two main components, an added white Gaussian component due to background photons and a component inversely proportional to some function of frequency.

The latter noise component, in systems with visual displays, is particularly objectionable because of the high frequency effect it causes in the direction normal to the raster lines due to imbalance between lines. It is reduced by 'D.C. restoration' where the detectors are periodically shown a reference radiation source and the input to the display simultaneously biased to a defined level and the bias clamped. Where an effective approach of this sort is possible the main residual inter-line imbalance may be due to the white noise component at the time the bias is determined and this can be minimised by one of the following methods.

The low frequency noise component can be reduced further by the 'overscan' method discussed in the next section, where each line is effectively sampled a number of times.

Another approach to reduce both forms of noise is to replace each pixel brightness by some weighted average of those of surrounding pixels. However, this blurs the image. It raises the question of finding the optimum compromise between resolution and noise for a given task.

For the visual recognition task the Company has found that, over a wide representative range of input signal to noise ratio, deblurring at the expense of increased noise does no harm and often improves performance. This is illustrated in the section under 'Aperture Correction'.

An alternative method is to replace each pixel by the median rather than the mean of its neighbourhood. This does not blur edges or ramps, provided there is only one edge in the neighbourhood at any instant, but smooths noise rather as the mean does.

Because this median filter is non-linear various forms are possible. For example several passes with the same filter leave parts unchanged, which a single pass does not change & progressively smooths the rest. A cross shaped window can be better than a square.

The median filter suppresses impulses very effectively and so, because it leaves ramps and edges unchanged it can be used to enhance compact sources against, for example, cloud background, very effectively by subtracting the filtered image from the original.

There is a variety of non-linear noise cheating algorithms which use logic to decide whether a given pixel brightness is likely to be due to noise and if so replace it by a better fit.

A method which evaluation to date suggests to be nearly as effective as a matched filter without the matched filter's need to know the target form, (and this is usually not known in practice) is to use a series of filters, localised in both space and frequency, but of different scales. A non-linear method analogous to the median filter can be used to remove the important noise from these localised filter outputs without materially affecting the important signal content. It is then possible to sum the filtered outputs and so reconstruct the scene with noise greatly reduced but with resolution minimally impaired.

The different scales of the filters give additional spin off in approaches to range invariance and the ability to allow the output of the channels at one scale to guide the processing of the outputs of other scale channels.

In automatic recognition and navigation systems, outline may be more invariant than contrast and so may form a simpler recognition criterion.

A first move in defining outline may be preprocessing to reduce noise as discussed above. This may be followed as a second move by determining the best elemental straight line edge in the neighbourhood of each pixel, by correlation with masks at different orientations and choice of that orientation giving the greatest output. This effectively integrates along the edge and further improves signal to noise ratio, provided the mask is not too long. If it is too long, edge curvature reduces signal to noise ratio, so that there is skill in the choice of mask.

The third move is likely to be to link these elemental straight edges to form curves using logic to decide which are mutually consistent and so valid and which are inconsistent and so due to noise. The process further enhances signal to noise.

Because there are three stages in the above process, it is necessary in assessing an approach, to include all 3 stages. For instance some preprocessors make elemental edge detection less effective and some elemental edge detectors do not provide good orientation information, helpful to the linking stage.

Usually the Signal Transfer Function of imagers is effectively linear, and for various reasons this may not be optimum.

One is that the eyes' response is not linear but more sensitive near black. Consequently it may be better to reduce image gain near black and increase it near white to pack more information into the region the eye can usefully use.

Sometimes the objects in a scene can be divided into two or more classes, perhaps natural components near ambient and internally heated components such as houses and people with very little scene content between the classes. An approach optimising this situation is histogram equalisation where the Signal Transfer Function is made to follow the form of the cumulative histogram of scene brightness. This allows the most used brightness region in the scene to use the greatest proportion of the grey scale.

Histogram equalisation can however, make an image more difficult to use. For example a histogram equalised chess board loses contrast between squares but amplifies noise within squares.

In some cases it is possible to improve performance by splitting the image into smaller sections each of which is homogenous in its basic characteristics and histogram equalise them individually. An important aim is often to enhance interesting objects and suppress the rest to minimise the observer's confusion. Local histogram equalisation however may enhance everything.

An analogous approach is to make the local areas 'Contrast' and 'Brightness' a function of, say local area standard deviation and mean level. Taken to the extreme this can turn the image into uniform noise and choice of optimum control signal and control function is important.

Information within the tank has been pulled out but so has a lot of the background information. As the standard deviation increases the image begins to look more and more noisy and some optimisation as to the choice of these parameters is needed.

The use of pseudocolour as well as brightness variation has the potential to vastly increase the information content of a display. It is important to connect colouration to human experience or the result just confuses. Blue = cold to red = hot, through the spectrum can have this property. Not only does the red suggest heat but the sky can be blue and ground green to yellow. Implementation of such a system may require the imager to sense absolute rather than relative temperatures.

In such systems it is necessary to arrange brightness as well as hue changes at edges for perception to use them effectively and there are various schemes to do this.

2.1 Overscanning

In most scanning systems it is usual for the raster pitch to equal the size of the detector. This results in a degradation of resolution in a direction normal to the raster and has serious effects on the range capability of the system.

Overscanning, in which the pitch of the raster is some fraction of the detector width, overcomes this problem and produces an isotropic image whose resolution is determined solely by the system transfer function. The idea of overscanning follows directly from the Nyquist sampling theorem. Normally a series of discrete samples are taken along the direction of the raster in accordance with this theorem but not in the direction normal to it. At least 2 samples per detector width are required in this direction. Overscanning will also reduce aliasing effects and the introduction of steps in otherwise straight edges. Results showing the effects of overscanning are shown in Figure 1 for 3 different ranges. The slide is a rather poor reproduction but it can be seen that overscanning produces a marked improvement in image quality. Imagery at a range of 1.5 km with overscanning is similar in interpretability to that at 0.75 km without overscanning thus doubling the range capability.

Overscanning is also desirable where it is necessary to register successive frames of data in order to track objects. It is possible that a point target which had previously completely covered 1 detector would, in the next frame lie between 2 detectors and therefore these pixels would only have half the energy. Comparison of pixels therefore when registering would become difficult. With overscanning at least one pixel would have the full information.

2.2 Aperture Correction

A non-recursive spatial filter for the removal of system blur was produced. This consisted of a 3 x 3 weighted matrix designed to collect energy falling on adjacent detectors, due to the spread of the system optics, and to add this back into the central detector.

Figure 2 shows the way in which this was done. 3 adjacent detectors are labelled A, B and C in this diagram. The freq. response of the centre element, B, is shown by the full curve and that of the function $\frac{B - A + C}{2}$ is shown by the broken curve.

Weighting this function by some factor k and adding to B results in an MTF with a better high freq. response. Trial and error have shown that a value of K of 0.3 is optimum.

Figure 3 shows the results of this 'aperture correction'. A sharpening effect is noticeable making interpretation easier.

Previous workers in aperture correction do not appear to have shown any useful improvement in recognition and it is believed that this was because raster overscan had not been employed.

Figure 4 shows the effects of aperture correction on noisy imagery. Improvement is still noticeable.

3. ENHANCEMENT OF SPATIAL RESOLUTION

Usually the Modulation Transfer function of the system can be measured. If so it is possible to apply an inverse filter to correct for it. Obviously at frequencies where the transfer function is so poor that all signal is lost in noise the pure inverse filter does more harm than good and some alternative strategy is needed for these regions.

A particular case which the Company has studied is that where the Modulation Transfer function is defined largely by the finite size of the detector element.

This may be true, for example, of an optimised staring matrix array detector. In the early development stage of such detectors it may be difficult to produce a good yield of large arrays. Perhaps the diameter of the array may be limited to a few tens of element. In this case there may be conflict between field of view needed and resolution needed. The best compromise may be to meet the field of view requirement and compensate detector blurring by a modified inverse filter.

The Modulation Transfer function of a square detector is a two dimensional sinc function, $\frac{(\sin x)}{x}$, an oscillation about zero dying away slowly.

Simulation has shown that, in this case, and making a reasonable input signal to noise ratio assumption resolution can be more than doubled.

Figure 5 shows a typical MTF for a 6X overscanned image and Figure 6 shows the corresponding inverse filter.

The bottom image in Figure 7 shows the 6X overscanned image used for the filtering work and Figure 8 shows the results of inverse filtering.

Noise at 43 cycles per picture width occurs on all these images due to a slight mismatch between the zeros of the MTF and the inverse filter.

(a)   is the result of applying the inverse filter in Figure 6. The next 2 images (b) and (c) have the maximum gain of the filter limited to ±20X and this has had little effect on the image.

In order to see how much information is contained in the higher frequencies successive lobes of the inverse filter were removed.

(d)   shows the effect of removing the information beyond the 3rd zero crossing

(e)   information beyond the 2nd zero is removed and

(f)   information beyond the 1st zero is removed.

It can be seen that useful information is available beyond the 1st zero of the MTF and it is important to retain this information -- beyond the 2nd zero however little is gained by its retention.

Image (h) shows the importance of retaining phase information. Here the phase of the 2nd lobe has been made positive.

Image (g) just shows the effects of tapering the 2nd lobe of the inverse filter to zero at 40 c.p.w. to cut out the 43 c.p.w. noise.

A sharp drop to zero in the filter function where the MTF falls to zero produces ringing in the image. The Wiener filter tries to overcome this by automatically reducing the gain in those regions where the signal falls towards the noise level. This filter produces poorer resolution than the inverse filter but also reduces ringing. Some kind of compromise between the two is desirable. This can be achieved by using weighted combinations of Wiener and Inverse filter. When α is equal to a half in this equation we have what is known as the Geometric mean filter.

$$\text{WIENER FILTER} = \cfrac{1}{H(f) + \cfrac{[Yn(f)]^2}{[Ys(f)]^2 \cdot H(f)}} \qquad (3.1)$$

where $H(f) = \text{Sinc}(\pi/fc)$, fc = freq. of 1st zero of system MFT

$[Yn(f)]^2$ = image noise power spectrum

$[Ys(f)]$ = image signal power spectrum

$(\text{Inverse})^\alpha \cdot (\text{Wiener})^{1-\alpha}$ \qquad (3.2)

Another approach which we are presently looking at is not to set the gain to zero in the regions of the MTF zeros but to insert whatever freq. and phase component gives the smoothest image or one having the noisiness expected at these freqs. from the known system noise characteristics.

When narrow bands of the frequency spectrum are set to zero it produces a ringing effect visible to the eye as such. Thus it is possible to determine the phase and amplitude necessary in the blanked region to suppress the ringing and this may be a good guess at the missing component.

One way to implement this is to choose the maximum entropy waveform and the classical approaches to this also ensure positive brightness solutions.

The prior knowledge that brightness must be positive produces a large improvement in resolution, where it can be applied, over linear restoration systems which if one attempts to push them too far produce overshoots instead of a narrower central peak to the corrected system spread function.

In general it does not apply to a.c. coupled thermal imagers. Nor does it help improvement of resolution of objects against positive rather than zero background.

However, in the final phase of command guidance, for example, and it is only in the final phase that enhanced resolution may be needed, missile and target are close together against what may be effectively uniform background and the system can adjust its d.c. restoration to make the background an artificial zero level. In these circumstances a resolution of four times better than the diffraction limit has been claimed[1].

In developing methods of target enhancement with respect to background a study was carried out on the spatial frequency characteristics of typical targets and backgrounds.

A characteristic mottling effect was noticeable in the spatial spectra of vehicles which were noticeably absent in those of backgrounds. This is illustrated in the following Figures which show the modulus of the Fourier transform and the logarithm of the modulus of the Fourier transform of typical vehicles and backgrounds.

Figure 9. This is the transform of a thermal image of a vehicle. Notice the mottling effect.

Figure 10 shows the transform of another vehicle. Again notice the patterns in the central lobe.

Figure 11 shows the transform of a typical background scene, plains. No characteristic mottling is present, the transform appears completely random.

Attempts to isolate these periodic mottlings were made by Fourier transforming these transforms again. These are shown in Figure 12. The images on the left hand side show the modulus of the Fourier transforms of the objects and those on the right the Fourier transform of the central part of those on the left. The top two pairs are of vehicles and show marked similarities whereas the bottom pair, are of typical background and are quite different. (The size of the lower right hand image was due to a horizontal scale change in the output program when this image was produced. All the other images were produced at an earlier stage).

So far no positive conclusions have been made from this study.

4. REFERENCES

1. Frieden, B.R.   *Picture Processing & Digital Filtering* Chapter 5, Edited by T.S.Huang, Springer-Verlag.

18-6

NO OVERSCAN

X2 OVERSCAN

X3 OVERSCAN

RANGE
KILOMETRES

0·75

1·0

1·5

Fig.1 Effect of raster overscan with diamond shaped analysing aperture constant system magnification

Detector array

A | B | C

RESPONSE OF EYE AT MAGNIFICATION PREFERRED FOR INTERPRETATION

$B + K \left(B - \dfrac{A + C}{2}\right)$

$B - \dfrac{A + C}{2}$

B

MOST IMPORTANT REGION FOR LONG RANGE RECOGNITION

MODULATION TRANSFER
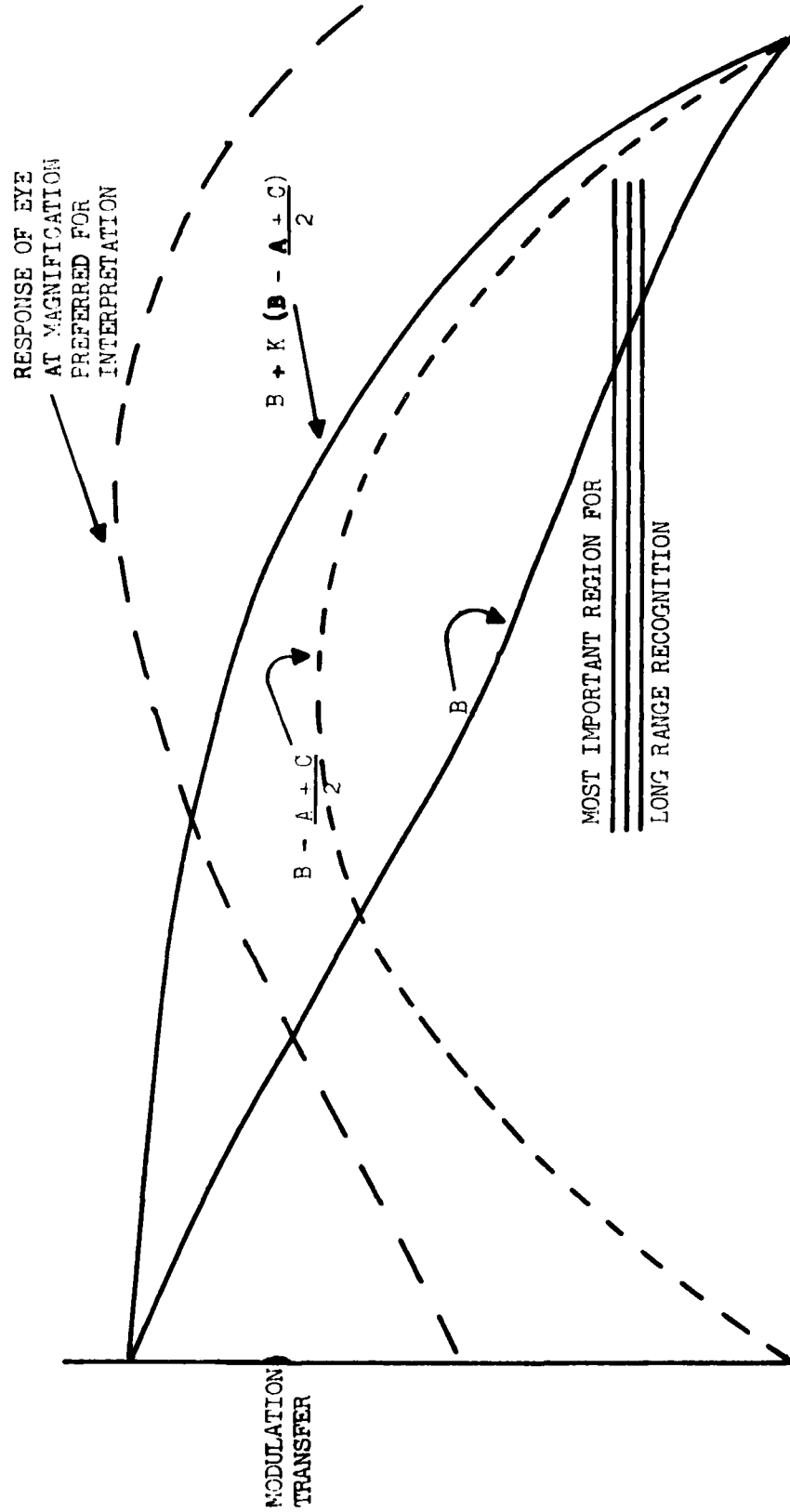
SPATIAL FREQUENCY

Figure 2

18-8

TREBLE OVERSCAN

NO NOISE

DOUBLE OVERSCAN

APERTURE CORRECTED

THE RIGHT HAND IMAGE OF EACH PAIR IS APERTURE CORRECTED

NO OVERSCAN

RANGE
(km)

0·75

·0

·5

Fig.3   Effect of aperture correction.

18-9

NO OVERSCAN   DOUBLE OVERSCAN   TREBLE OVERSCAN

RANGE (km)   0.75   1.0   1.5
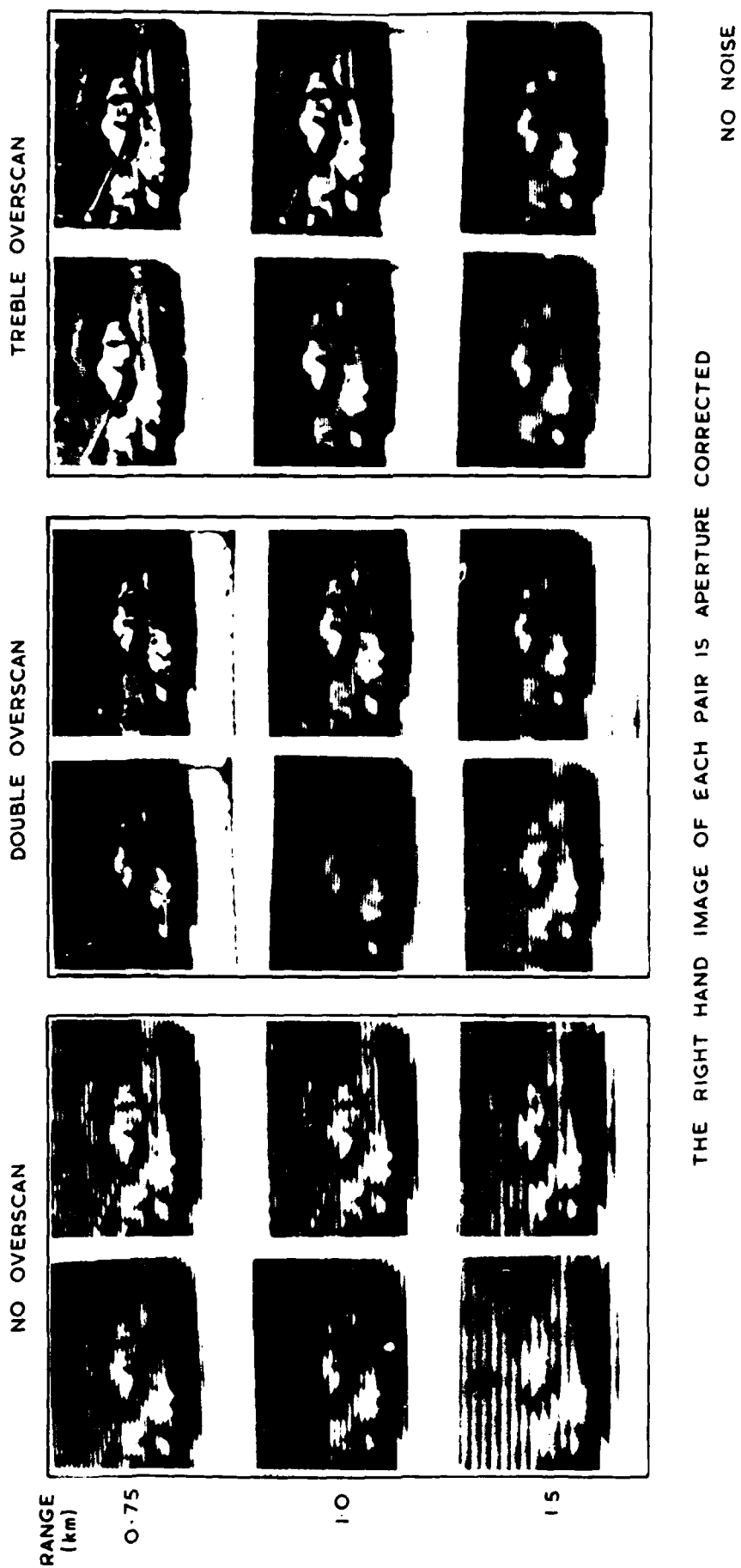
THE RIGHT HAND IMAGE OF EACH PAIR IS APERTURE CORRECTED

NOISE STANDARD DEVIATION $\frac{1}{30}$ PEAK WHITE

Fig.4   Effect of aperture correction
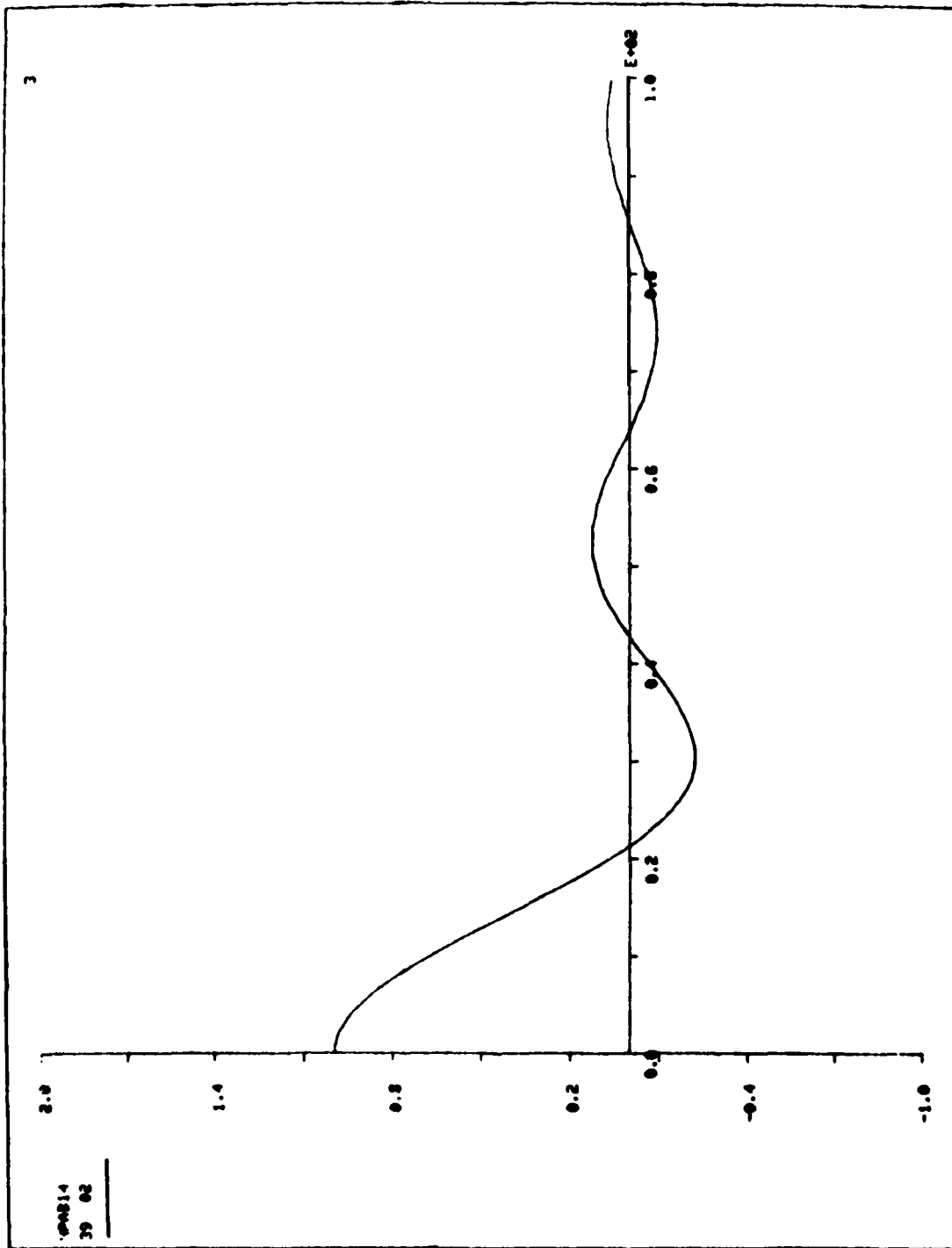
Figure 5
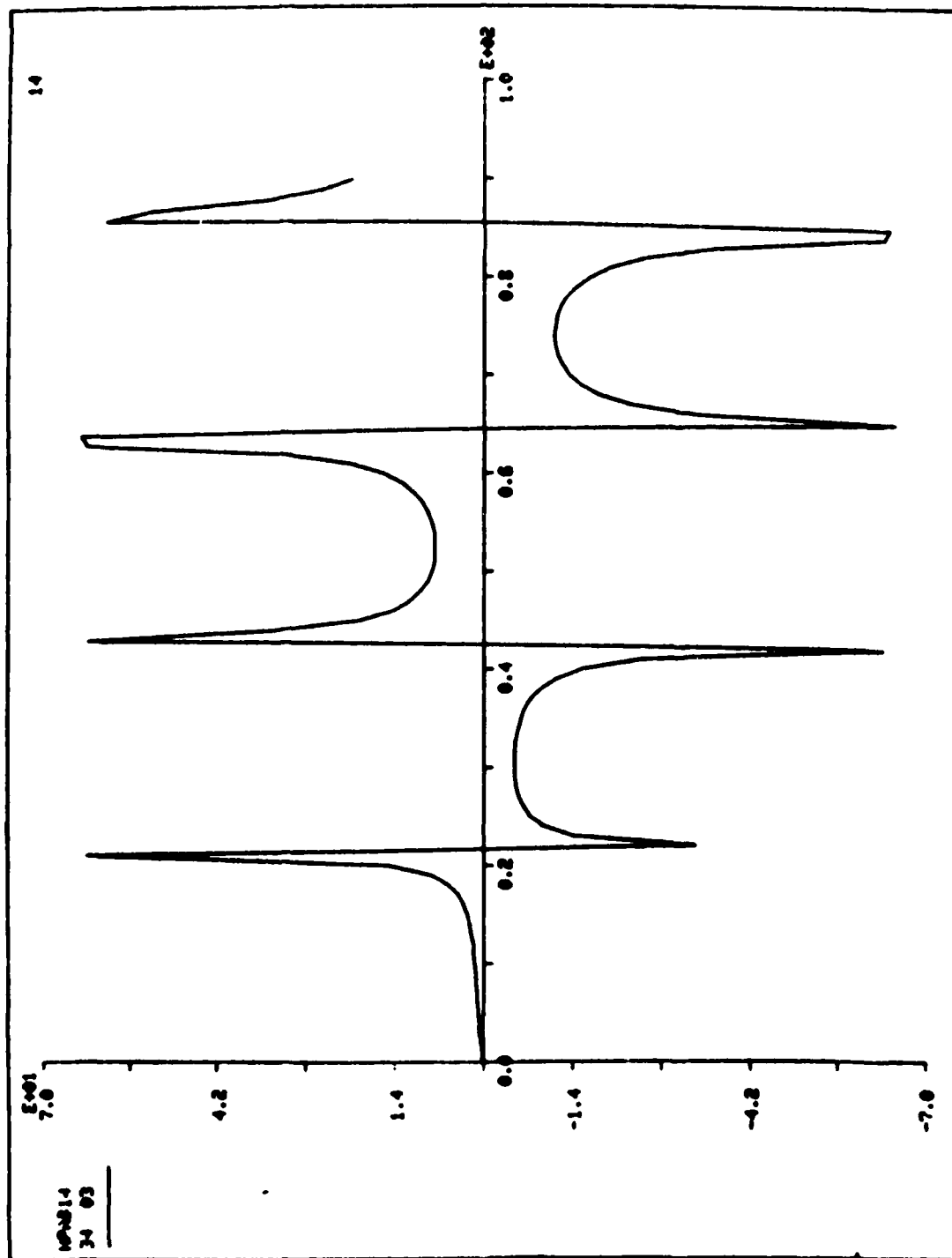
Figure 6
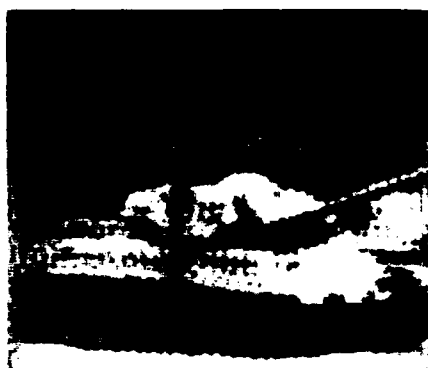
18-11

Figure 7

Figure 8

CYCLES PER
PICTURE WIDTH

(d) Logarithm of transform

(c) Modulus of transform

Figure 9

CYCLES PER
PICTURE WIDTH

(d) Logarithm of transform

(c) Modulus of transform

Figure 10

CYCLES PER PICTURE WIDTH

(d) Logarithm of transform

(c) Modulus of transform

Figure 11

Figure 12

IMAGE ENHANCEMENT IN REAL TIME
by
H. Yndestad
Norwegian Defence Research Establishment
Po Box 25
2007 Kjeller
Norway

SUMMARY

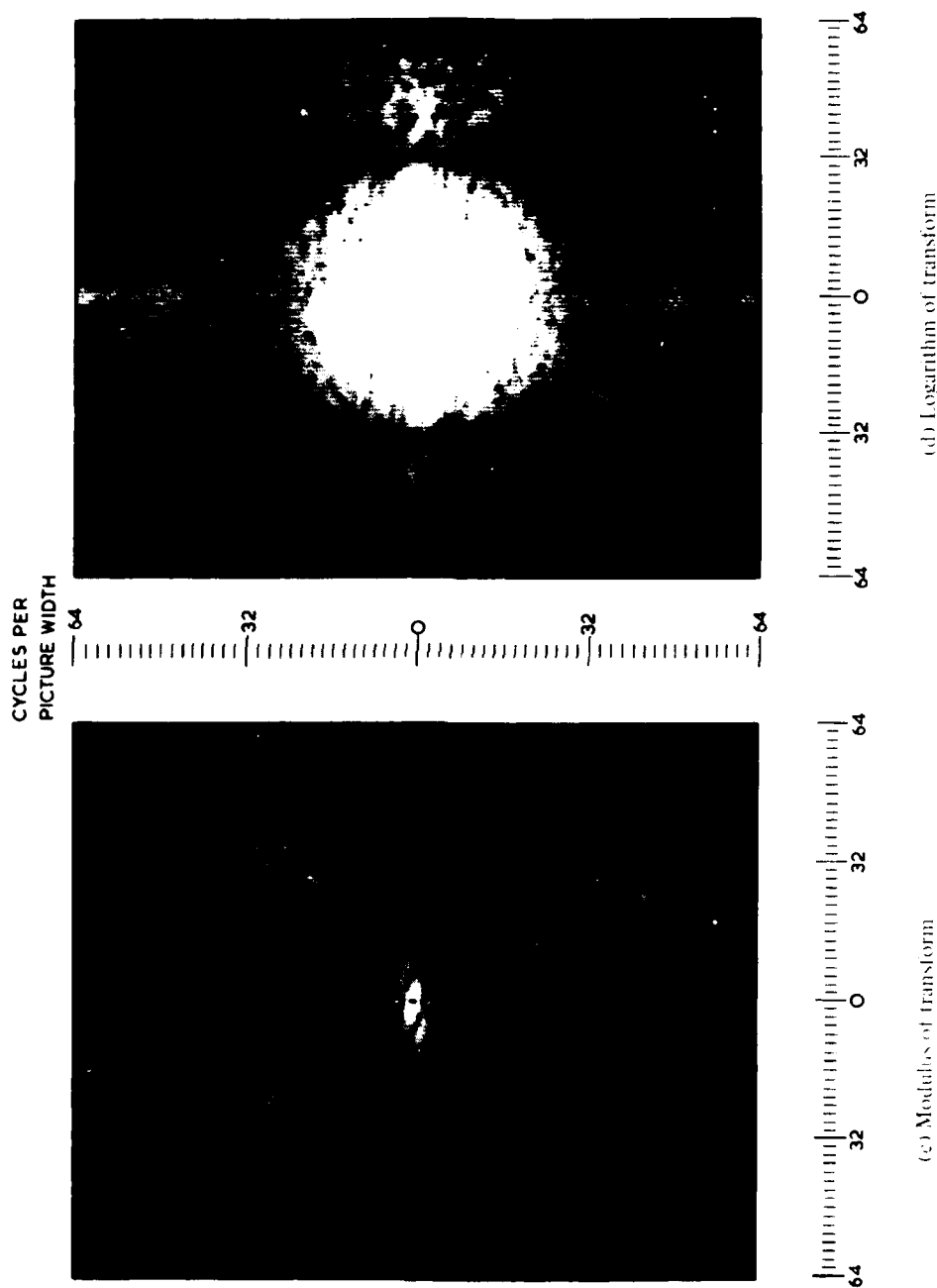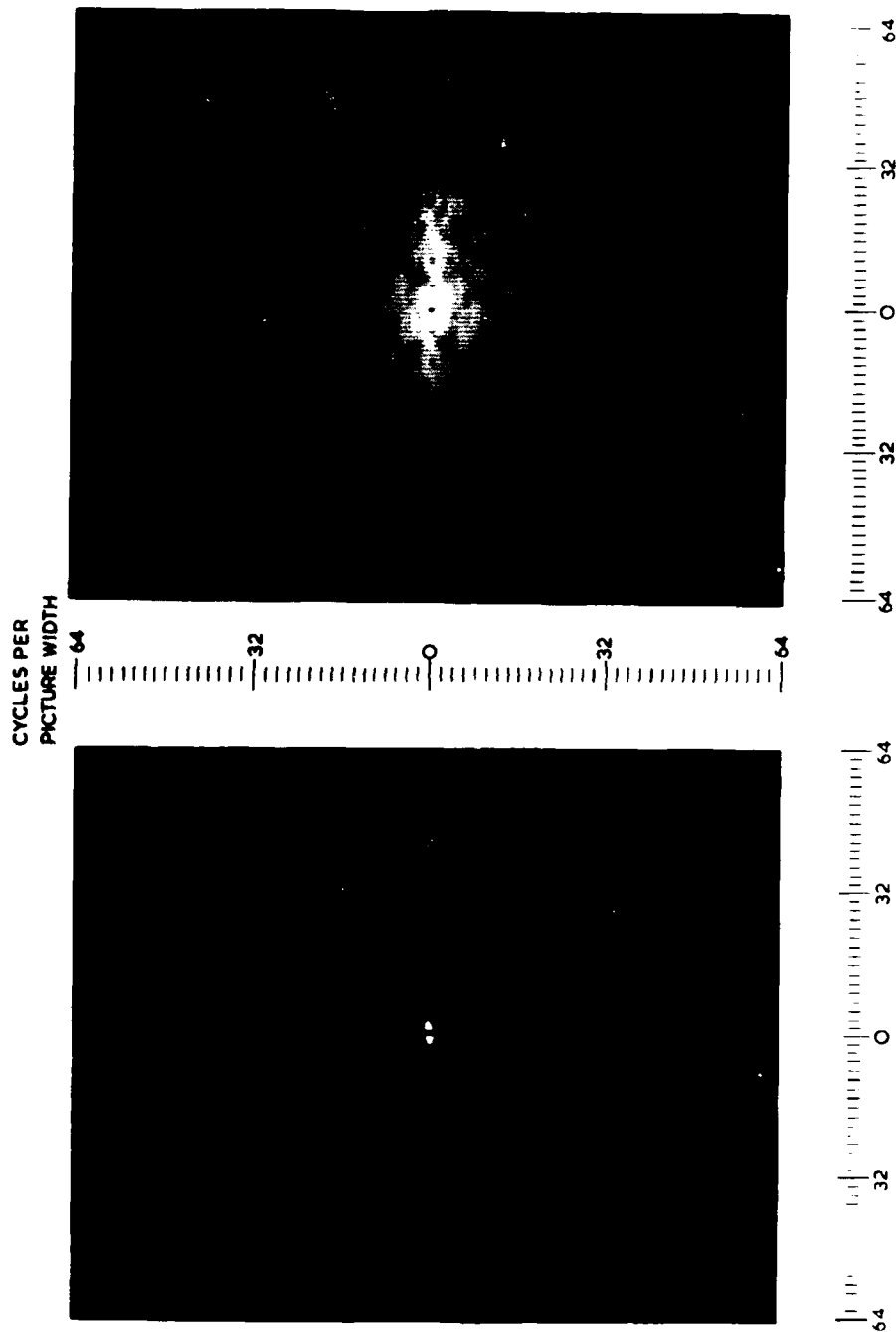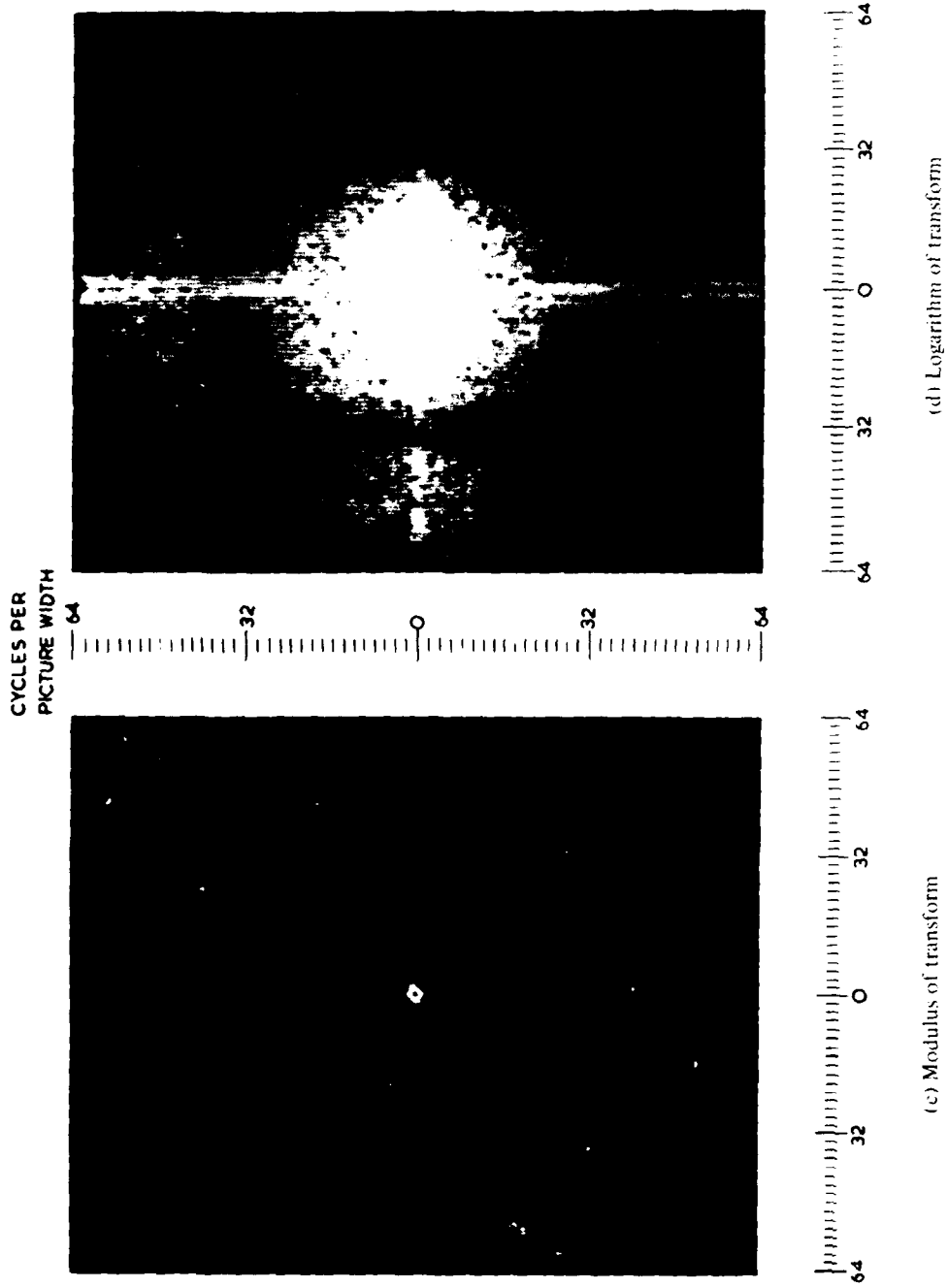This paper presents an algorithm and a planned computer architecture for image enhancement in real time. The algorithm for image enhancement will reduce dynamic range and enhance contrast in an image by homomorphic filtering.

The filter function is separable in column- and in line-direction, and the filter is implemented as a linear phase frequency sampling filter. Two-dimensional filtering can then be achieved by filtering a picture first line by line and then column by column in a recursive filter. Dynamic range reduction and contrast enhancement are controled by two parameters. The algorithm needs 32 arithmetic operations.

The image processor is a set of processing elements, each having a specific task. One such task can be the described algorithm for homomorphic filtering. The processing element is built up by a micro-processor for local control, a bit-slice processor for data transfer control and a specialized arithmetic unit for signal processing. Data transfer between processing elements is carried out in a high speed ring network.

1 INTRODUCTION

The illumination on a scene often varies strongly. The dynamic range of imaging sensors can be about 50 dB or more. A display device such as a cathode-ray tybe typically has a dynamic range less than 30 dB. When a video signal from a wide dynamic range sensor is to be displayed on a display device, some form of processing is therefore required to minimize the loss of information.

Spatial homomorphic filtering is a method that has good dynamic range reduction and contrast enhancement properties with a minimum loss of information [1]. Because of the good properties of this method, it has been of interest to design an image processor that can do the necesary processing in real time.

Spatial filtering in real time requires an image processor with a very high processing capacity. It is therefore important to have a filter algorithm which is easy to implement, and the image processor must have an architecture that gives flexibility and a high processing capacity.

2 DEFINITIONS

An illumination component is a digital representation of the illumination on a scene. A reflectance component is a digital representation of the reflectance at the scene.

A picture is a two-dimentional array of numbers that is a sampled and digized representation of the luminance from a scene. A picture element is an element of this array.

3 HOMOMORPHIC FILTERING

The picture elements can be written [1] as

$$x(n,m) = x_i(n,m) \cdot x_r(n,m) \qquad (1)$$

where n and m are the discrete spatial variables.

The illumination component $x_i(n,m)$ is always positive and has a wide dynamic range when there are shadows in the scene. In most cases this component has most of its energy at the lowest spatial frequencies in the picture [1].

The reflectance component $x_r(n,m)$ is always above zero and below one. This component modulates the contrast in the picture and it is spatially more high frequent than the illumination component. This gives the possibility of a dynamic range reduction and at the same time a contrast enhancement, by some kind of high pass filtering.

There are some difficulties with a direct linear filtering of the picture. A linear filtering will introduce both negative and positive element values. A second problem is that the illumination component $x_i(n,m)$ and the reflectance component are multiplied. It can be shown [2] that it is easier to filter them independently when they are added.

Both of these problems can be solved by taking a logarithmic transform of the picture element values before the linear filtering. This gives a new set of components which are additive.

$$g(n,m) = \ln[x(n,m)] = \ln[x_i(n,m)] + \ln[x_r(n,m)] \qquad (2)$$

If a linear filter modifies the illumination component by a factor B and the reflectance component by a factor A, then the filtered picture elements in additive form will be

$$p(n,m) = B \cdot \ln[x_i(n,m)] + A \cdot \ln[x_r(n,m)] \qquad (3)$$

The modified components are then transformed back to multiplicative form by an anti-logarithmic transform

$$y(n,m) = \exp[p(n,m)] = [x_i(n,m)]^B \cdot [x_r(n,m)]^A \qquad (4)$$

The dynamic range is reduced when B < 1, and contrast is increased when A > 1. This type of filtering is called homomorphic filtering.

3.1 The linear filter

A simple way to make the linear filter is to subtract spatially low-pass filtered elements from the unfiltered elements as illustrated in figure 1. This can be expressed as

$$p(n,m) = A \cdot (\ln[x_i(n,m)] + \ln[x_r(n,m)]) - (A-B) \cdot (\ln[x_i(n,m)] + \ln[x_r(n,m)]) * h(n,m) \qquad (5)$$

where A and B are constants, * is the convolution operator and h(n,m) is the impulse response of the low pass filter.

If there is no overlap between the spatial frequences of the illumination and the reflectance component, a low pass filter h(n,m) can be found that will suppress the reflectance component. Then (5) can be modified to

$$p(n,m) = B \cdot \ln[x_i(n,m)] + A \cdot \ln[x_r(n,m)] \qquad (6)$$

which is identical to (3).

In a practical filtering there will be some overlap between the spatial frequences of the two components. This will limit the possibility of increasing the contrast in the picture.

3.2 The low-pass filter

A useful filter function for the low-pass filter is a normalized Gaussian.

$$H(k,l) = \exp(-(k^2 + l^2)1/2 \; \sigma^2) \qquad (7)$$

where k and l are the discrete spatial variables and σ is a selected constant.

This filter function has two important properties. It is separable in line- and column-direction. Twodimensional filtering can then be achieved by filtering the picture first line by line and then column by column. This is of importance in a real time processing, as the number of arithmetic operation is much reduced compared to a circular convolution. The second important property by this function is that it is rotationally invariant.

The low-pass filter can be implemented as a frequency sampling filter [3]. This filter is easy to implement and to approximate to a Gaussian filter function. The bandwidth is tuneable by one parameter and it can be implemented to have exact linear phase.

The difference equation for a simple frequency sampling filter can be written as

$$p(n) = a(n) + b(n) \qquad (8)$$

where

$$a(n) = r \cdot a(n-1) + (g(n) - r^N \cdot g(n-N)) / N$$
$$b(n) = 2 \cdot r \cdot \cos(2\pi/N) \cdot b(n-1) - r^2 \cdot b(n-2) - (g(n) - r^N \cdot g(n-N)) \cdot (g(n) - r \cdot g(n-1)) \cdot \cos(\pi/N) \cdot |H(1)|/N$$

where

g(n) is the input signal

|H(1)| is a sample of the normalized Gaussian.

N is the number of samples of the impulse response. This parameter controls the bandwith of the filter.

r is a constant less than one.

The linear phase property of (8) will introduse a delay of N/2 samples. The equation (5) will then be modified to

$$p(n,m) = A \cdot (\ln[x_i(n-(N-1)/2, m-(N-1)/2)] + \ln[x_r(n-(N-1)/2, \; m-(N-1)/2)]$$

$$- (A-B) \cdot (\ln[x_i(n,m)] + \ln[x_r(n,m)]) * h(n,m) \qquad (9)$$

where N is an odd number.

This filter algorithm needs 32 arithmetic operation per picture element. It has three parameteters. In (9) B will control dynamic range reduction, A controls contrast enhancement and in (8) N will control the bandwidth of the low-pass filter.

4 THE IMAGE PROCESSOR

A general purpose computer does not have the necessary compuing capacity for image processing in real time. A multiprocessor configuration will have the computing power, but the practical implimentation will be complex.

To reduce volume, costs and power consumtion, the image processor has to be specialized. A specialized computer architecture will lose flexibility. It is therefore important to have an architecture that is modular.

In a modular computer architecture it is necessary to have a powerful data communication system. Such a computer architecture is shown in figure 2. It is build up by a number of processing elements (PE) connected to a ring network. A processing element is an intelligent unit that is specialized to a specific task. Such a task can be a video digitzer, noise cleaning, homomophic filtering or a display processor. It can also be a general purpose computer or a backend database system.

The ring network is a high speed packed-switched transmission system for carrying data between processing elements. Compared to a more traditional bus system, the ring network has two significant advantages. It requires no global control and it has a higher data transfer capacity. More than one processing element can simultaneously put a data word into the ring. The ring will then transmit the data word to the receiving process elements.

4.1 The node

A node is an element of the ring network which transmits a data word between processing elements. The node has a register, an input queue and an output queue as shown in figure 3. The register carries a data word by receiving it from one node and sending it to the next. The input queue receives data words from a processing element and puts it into the register, and the output queue receives a data word from the register and put it into a processing element. A queue can be implemented by a First In-First Out Memory.

The data word consists of an information word, a sender's adress, a receiver's adress and one bit that indicates if the word is filled or empty.

When a data word is transmitted, a processing element put it into the queue. The first data word in the queue is put into the register when the node receives an empty word. Then it is carried around the ring. When it arrives to the adressed node, it is put into the output queue and the data word at the ring is marked as empty. The adressed processing element will then fetch the data word from the output queue of the node.

4.2 The processing element

A processing element (PE) is an intelligent unit that can either be a general purpose computer, a backend database processor or a specialized processor for a specific task.

If the PE is an processor that executes homomorphic filtering in real time, it can be organized by a micro-processor, a bit-slice processor and a specialized arithmetic unit as shown in figure 4.

The micro-processor controls an PE. It will be able to communicate to an other PE, compute filter parameters and control the bit-slice processor.

The bit-slice processor is a high-speed data transfer controler. It controls data transfer within the specialized arithmetic unit and the data transfer to and from the node.

All the signal processing can be done in a specialized arithmetic unit, which is directly wired according to the filter equation. For each multiplication in the filter equation, there is a monolithic multiplir and for each addition there are adder components. Registers can be introduced at suitable locations to achieve pipelineing. The logarithmic and anti-logarithmic transform can easily be obtained by look-up tables in Read Only Memories. The picture is then filtered, simply by clocking the picture element values at clock rate speed through the arithmetic unit, first line by line and then coloumn by coloumn.

In some cases it may be possible to implement the specialized arithmetic unit by a custom designed charge coupled device or some other analog processing device.

The computer architecture gives an image processor that has a relatively small volume and power consumtion. The lack of a full programable flexibility is compensated by its modularity. A new PE can easily be plugged into the system.

19-4

References:

[1] Oppenheim, A V   — "Nonlinear Filtering of Multiplied and Convolved Signals",
    Schafer, R W   Proceedings IEEE (August 1968)
    Stockham, T G

[2] Boulter, J F   — "Use of Spatial Filtering to match Wide Dynamic Range Gray-
                               scale Imaginary to a lower Resolution Display", Defence
                               Research Establishment, Valcartier, Québec, Canada,
                               (January 1977)

[3] Rabiner, L R   — "Recursive and Nonrecursive Realizations of Digital Filters
    Schafer, R W   Designed by Frequency Sampling Techniques", IEEE Trans Audio
                               Electroconst, Vol AU-19, No 3, (September 1971)

[4] Yndestad, H   — "Sanntids bildeprosessor", TN-E-1021, Norwegian Defence
                               Research Establishment, PO-Box 25, 2007 Kjeller, Norway
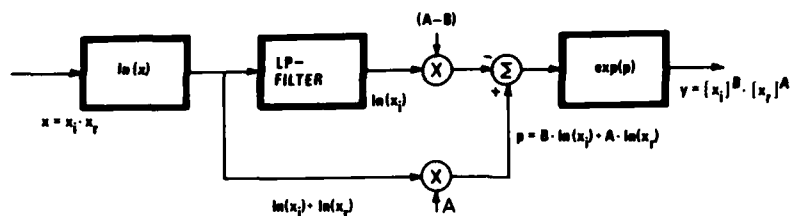   (January 1979)
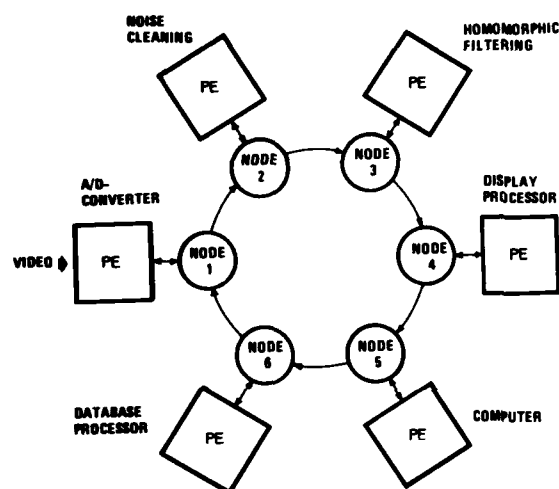
Figure 1 Algorithm for homomorphic filtering
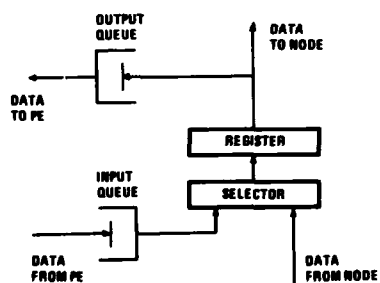

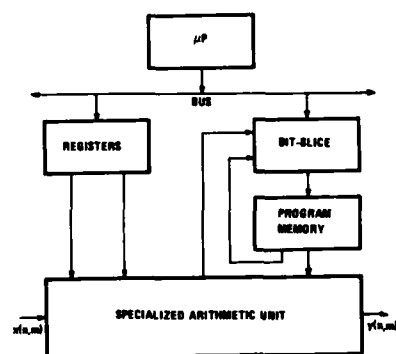Figure 2 Image processor architecture

Figure 3 A node

Figure 4 Processing element for homomorphic filtering

SEMI-AUTOMATIC AND AUTOMATIC EXTRACTION OF OBJECTS FROM AERIAL IMAGES

W. Kestner, M. Sties
Forschungsinstitut für Informationsverarbeitung
und Mustererkennung (FIM/FGAN)
Breslauer Straße 48, 7500 Karlsruhe
Fed. Rep. Germany

SUMMARY

Different methods for the extraction of objects from aerial images are presented. Unlike other methods which process the complete image systematically, we developed object guided methods, which are applied only to those parts of the image where objects or parts of objects have already been detected, i.e. where a continuation of an object is probable.

Basic principles as well as details of the extraction methods are explained. The methods differ with respect to the local precision of the results, the applicability to different object types and the required image quality. Local operators are described which evaluate grev level diagrams in order to detect object continuations.

The methods have been implemented on a computer DEC PDP 11/70. Results are presented.

1.   INTRODUCTION

Due to the variability of objects and the complexity of object interrelations, the interpretation of aerial images is almost completely reserved to specially trained human interpreters. This is true for each of the different fields of application. An automatic system to solve this interpretation task is not known to date (Kyoto, 1978). In this context the term 'image interpretation' stands for the highest level operation, which comprises any other operation such as image segmentation, object detection, object classification, evaluation of object contours, object attributes or object interrelations, etc.

For special subproblems of the interpretation task such as object classification without explicit definition of object outlines, several automatic or at least semi-automatic methods have been developed (DFVLR, 1980; Ouiel, F., 1979). Other problems related to the detection of line and region objects in aerial imagery are investigated by a growing number of scientists (Zucker, S.W., 1977; Keng, J., 1977; Montoto, L., 1977; Braconne, S., 1978; Nagao, M., 1978; Nevatia, R., 1978). The systematic processing of the image matrix, which is a basic characteristic of all methods cited, tends to require much computing time and sometimes leads to noticeable error rates, due to the variability of the objects.

The object extraction methods to be described differ from the above mentioned methods in a number of basic characteristics (Compare also Quam, L.H., 1978).

2.   BASIC PROCEEDING

In this paper object extraction is defined as the compilation of coordinates of the object outline (contour) or object centre line on the basis of a step by step tracing, i.e. a continuation of the contour or centre line is evaluated by prediction and verification of features and locations. This procedure requires identification of at least one point of the contour or centre line. The problem of how to identify such a point either visually or automatically (i.e. the initialization of the procedure) will be discussed in section 5.

The first characteristic of the extraction methods is their capability to adapt special object features or at least special features of different object types. Three different object types are distinguished which make up most of the objects recognizable in an aerial image:

   - line shaped objects are distinguished by their extreme length/width ratio.
     Roads, railroads, rivers, etc. belong to this object type. The result of the
     extraction of an object of this typ will usually be the object's centre line.

   - region like objects can be described as having a large enough area and a closed
     contour line to separate them from other objects. Examples of this object
     type are an urban area, a forest, an agricultural field, etc.

   - point like objects have only a short length and a small area such as single
     trees, houses, vehicles. The description of the location of objects of this
     type is usually confined to a pair of coordinates.

All three object types are treated differently either by applying different extraction methods or by inputting special feature values such as the width of a road or the specific texture of a forest. By the exploitation of this kind of a priori knowledge about certain object attributes, we succeeded in developing simple and fast extraction methods, which, at the same time, produce reliable results.

It is another basic principle, that the extraction methods do not process the complete image matrix systematically and uniformly, but work in limited parts of the image only. With the concept of a local initialization, it should obviously be more successfull to to analyze the very near neighbourhood of the actual location to detect a possible continuation of the object or object contour than applying a universal contour detection operation to the complete image matrix. This limitation of the area to be screened results in a considerable reduction of data to be processed. On the other hand we also get rid of the burden of searching for objects or object contours at locations within the image data where there is a very low probability to find one.

The basic algorithm of the extraction methods performs an analysis of a locally limited grey level function. As we can predict the location and the most probable orientation of further object parts to be detected, we even do not have to process a two-dimensional submatrix of grey level values, but can confine the algorithm to the analysis of one-dimensional grey level profiles which are perpendicular to the predicted object orientation, e.g., we are looking for the very next cross section or a limited sequence of continuing cross sections to be contained in the grey level profile at a predicted location. Details of the analysis of the grey level profile are discussed in section 4.

The possibility to improve t' local precision of the results step by step is to be mentioned asthe lastbasic pri ·ᵒⁱᵉ. To accommodate the differing requirements of local precision in the differentᵈᶠ ᵒf applications, several extraction procedures can be initiated sequentially, each which uses the resulting output of the preceeding procedure as input to be refineᵥ. With a growing amount of computer power a first rough approximation of an object's contour may be refined step by step to yield a precise approximation.

3.   SPECIALIZED EXTRACTION METHODS

3.1.   Extraction of line shaped objects

3.1.1.   Incremental method

The incremental method basically proceeds as follows:

- suppose that two points p_0 and p_4 of the centreline of a line object have already been found (see figure 1)

- connect p_0 and p_4 by a straight line

- predict the location of the next point p_2 of the object's centreline at a certain step width beyond p_4

- compute the coordinates of all points of a semi circle through p_2

- compile the grey levels of all of these points into a grey level diagram which shows a display of the cross section of the grey level situation at location p_2 (see figure 1b)

- analyze the grey level diagram to discover features of the line object

- compare the candidate features with the actual features of the points p_0 and p_4 which have been confirmed.

- in case of compliance decide on the true location of p_2, adapt the features and go on.

Locally guided by the results of previous steps this method proceeds incrementally along the object's centreline. All object points which have been detected and accepted by the algorithm are displayed immediately for visual control by the human operator. During the object extraction the procedure's parameters such as width of the object, average grey value of the object, average contrast with the neighbourhood, and step width are continuously updated to the actual values to adapt possible changes. The procedure has be completed by the grey level analysis of several concentric full circles around the actual location for two reasons mainly (Groch, 1980):

- to overcome local distortions of the object which have been detected by the fact, that no normal continuation of the object could be found while analysing the semi-circle as mentioned above

- to evaluate the grey level situation completely in the case of two or more line objects crossing each other which could have been detected by the fact the more than one normal continuation of the present object was found while analysing the semi-circle.

3.1.2.   Area slicing method

The area slicing method works in a predetermined limited part of the image, called "area of interest" which usually has a rectangular shape. The area of interest is expected to circumscribe a interesting part of the line object completely (see fig.2). At the same time the orientation of the rectangle is supposed to comply with the predominant orientation of the object. It has been found that a complete analysis of the 2 D grey level matrix of the area of interest is not necessary to discover the continua-

tion of a line shaped object, but that the analysis of 1 D sample lines which contain cross sections of the object, is sufficient. The sample lines are perpendicular to the main axis of the rectangle. The results of the analysis of the sample line's grey level functions are combined in order to form sets of collinearly connected points (Bausch et. al., 1979 a); at the end these sets of collinear points are linked to form lines.

3.2   Extraction of region like objects

3.2.1   The binarization method

Certain region objects which are characterized by a rather uniform overall grey level and a high contrast with their surroundings, such as a dark forest or a lake, may be extracted by binarization. For this purpose, an area of interest is circumscribed to the object. Within the area of interest a global evaluation of grey levels is applied to determine a threshold which will, most likely, seperate object pixels from non object pixels. Thus, binarizing of the area of interest results in several subsets of object pixels, non object pixels and neutral pixels of a reject area, which cannot be assigned without doubt neither to the object nor to the non object area in this processing step. The reject area may be subjected to a postprocessing algorithm which decides on the of a local majority against or in favour of the object pixel's subset. The last step consists of a merging of all object pixel subsets and the detection of its border.

3.2.2   The radius method

The method aims at the detection of a few sample points of the border line (contour) of the region object. The samples are connected by straight line segments and are considered to be a rough approximation of the true contour of the object. To detect the sample points, grey level functions are analyzed, which are assembled from clockwise arranged radii originating at a central point within the region object. The intersections of the radii with the contour are detected as characteristic changes of the grey level function (edge detection, see also section 4 and fig. 3a and b).

3.2.3   The modified area slicing method

To refine the rough approximation detected by the radius method a modified version of the area slicing method may be applied to each segment of the approximation polygon (Bausch, U. et. al., 1979 b). Again a number of grey level functions, perpendicular to the orientation of the respective polygon segment, is assembled, cross sections of the true contour which should be contained in the grey level functions, are predicted and verified via an edge detection technique and the local collinearity of adjacent sample points is evaluated (see fig. 4).

The modification of the method refers to the edge sensitivity of the grey level function analysis which contrasts with the line sensitivity of the same analysis within the original method; both are discussed in section 4.

3.3   Extraction of point like objects

One method has been developed to extract point objects. Remember that "point objects" must not consist of a single pixel; they should comprise a few pixels which cover a small, but noticeable area: they differ from "region objects" by their small size and their compact form. The method proceeds in several steps and works only in a limited area of the image. With the results of the line and region objects extraction we confine the search for specific point objects to promising areas such as:

  - vehicles are assumed to appear only on or near any kind of a road

  - houses and other buildings are expected to exist in urban areas or
    in the neighbourhood of roads only.

The process begins with the analysis of onedimensional grey level diagrams of two test lines, one perpendicular to the other (see fig. 5). If both diagrams contain features which indicate the existence of a point object considering its size, shape and orientation, a cue is assigned to that location and the verification is initiated. Verification consists of the analysis of the grey values of a twodimensional submatrix circumscribing the cue area. At first an attempt is made to seperate the supposed object from its neighbourhood by binarization technique. The threshold is derived from the previously successful diagram analysis. The resulting set of object pixels is subjected to a classification algorithm, with takes into account certain shape features such as length/width ratio, area/perimeter ratio, absolute size, etc.

4.   Analysis of 1 D grey level functions

The analysis of 1 D grey level functions, the size and the location of which can be determined by previous results of the extraction process, is a common feature of all object extraction methods mentioned above. The analysis algorithm is different for the different object types and is outlined as follows.

4.1   Grey level profile of a line object

The cross section of an undistorted line object (such as a clearly visible part of a road or a river) in a digitized aerial image shows a characteristic profile in a grey level function: a bright road in dark surroundings will result in a positive peak (location A in fig. 6 a); the dark water of a river which contrasts ideally to a bright surrounding, will lead to a negative peak of the function (location B in fig. 6 a)

In real world images the object profiles in grey level functions will not have such an ideal form, but will be degraded by various distortions. Thus, the analysis of a grey level function starts with the assumption that the profile of a line object is contained. To prove this assumption, a value C is computed for each pixel of the grey level function. The amount of C stands for the compliance of the actual limited grey level function at that location with the assumption. The value C is computed by arithmetic combination of the absolute grey level differences ΔG_1 and ΔG_2 with the "width" ΔW of the profile (see Fig. 6 b). The number of pixels to the left and to the right of the current location X is variable up to an upper limit in order to maximize the absolute value of ΔG_1 and ΔG_2 ; the upper limit of the variable test distance is determined via a priori knowledge about the type of object to be detected and the scale of the image. In fig. 7 an actual grey level function and the function of the profile assessment values C are shown. It will be noted, that the values C of the peak parts of the grey level function rank highest, whereas for uniform (i.e. featureless) parts of the grey level function the values C tend to be zero.

4.2   Grey level profile of region object contours

The cross section of a region object's contour shows, in an ideal case, characteristic features (see fig. 8). In real images such an ideal profile is often degraded by specific structures of the object, by variable contrasts and by noise.

To test again the assumption, that a cross section of a region objet's contour is contained in the grey level function, a modified value of compliance C_m is computed for each pixel of the grey level function. The value C_m results from an arithmetic combination of the following local features of the grey level function (see fig. 8):

- the grey level difference ΔG measured in section M of the function, where X is in the centre of M

- the lack of grey level differences E_L and E_R ("evenness") measured in the sections L and R of the function

- the difference of grey level variances measured in the sections L and R of the function.

The widths of the sections M, L and R is variable with an upper limit, which is determine by a priori knowledge about the objects to be extracted and the scale of the image. The actual width of M is chosen to maximize ΔG, whereas the actual width of L and R is chosen to minimize the respective grey level differences, i.e. maximize the evenness of the function within these sections.

The operation may be interpreted as superimposing a gauge of an ideal "step edge" of variable but limited size on each location of the actual grey level and computing a value of compliance between this gauge and the local function values. In fig. 9 an example of a grey level function of a real image and the computed values C_m are shown.

5.   Initialization of the methods

The initialization of the methods described consists of two main parts:

- input of feature values, parameter values, thresholds etc.

- definition of starting locations, from where a continuation of the object or the object contour can be traced

To solve the first problem a set of default values has been compiled in numerous test. These values have proven to be useful for a variety of applications and can be changed by operator interaction if necessary.

To solve the second problem, the most simple and most flexible, but not the fastest method consists of a visual identification of suitable starting locations by the operator with manual input of the respective coordinates. To support the operator at this task an automatic method has been developed which will be described for the example of roads.

A starting location for the road extraction method is defined as a clearly visible, unsdistorted, small section of a road. Such a section of a road is accepted only if it satisfies certain conditions which are tested sequentially. At first a grey level function analysis is initiated along several equally spaced horizontal and vertical image lines. Each location where the profile of a line object has been detected, is marked as a cue for a detailled investigation. This investigation tries to verify a cue by a full analysis of the grey level functions of two concentric circles around the cue

location. The cue is accepted as a starting location for road extraction only if the following conditions are satisfied:

- two and not more than two cross sections of a line object were detected at the inner circle

- two and only two cross sections of a line object were detected at the outer circle

- all cross sections detected and the cue itself are collinear

- the grey level variation of all pixels between the far most cross sections detected does not exceed a specified limit.

For each starting location which has been accepted, the location coordinates and its orientation are passed to the extraction method to enable the first prediction of the continuation (Groch, 1979).

6.   RESULTS

All methods have been implemented on a computer DEC PDP 11/70. Due to the limitations of the computer, all programs could not be run as a single task; the different tasks which have to be installed, interface with each other via data files on magnetic disk.

The methods were tested with aerial imagery of scales from 1 : 10000 to 1 : 100000. Subimages of a size of 6 cm x 6 cm were digitized using a DICOMED device. The spacial resolution was adapted to the scale of each image, that each pixel of the image matrix represented a circle area of 2 m to 5 m in diameter on the earth. The following examples are reproductions of displays from a COMTAL screen.

Figure 10 shows results of extracted roads with the incremental method. Only one starting location on the horizontal part of the road at the right side was provided and no interactive correction or postprocessing was necessary.

Figure 11 gives an example of the extraction of a variety of line objects. The result was produced by the combined application of the incremental and the area slicing method. Both methods complement each other very well: while the incremental procedure has advantages at curved lines and at the detection of all branches of an intersection, the area slicing method proves superiority at distorted sections, occluded parts and variable features of the objects.

In figure 12 an example of the extraction of a forest region is shown. A first approximation of the forest's contour was produced by the radius method and was used to guide the modified area slicing method. The resulting precise approximation of the contour still has a few gaps which could be filled by a postprocessing operation. The last example is shown in figure 13, which demonstrates the method for extraction of point objects. The display consists of roads and surrounding strips in an urban area. All the cues for vehicles which have been detected during the first step of the procedure are marked with a cross and those object which were finally accepted, are marked with small rectangles. Some imperfections of the performance regarding especially cars parked at the border of the road may need a refinement.

7.   CONCLUSIONS

In conclusion the following facts be stated as a support of the solutions explained above:

- interaction of a human operator which still proves to be necessary, has been concentrated to the initialization of the methods and the correction and completion of results

- due to the reduction of image data to be processed, fast computing can be realized. An assessment has shown, that with a modern computer system the extraction of a complete network of roads from an image of an average 2 km x 2 km section of a civilized country needs not more than 10 sec of processing time (without data input and output and without interaction)

- the results are reliable although not always complete. But the interactive completion proves to be easier than the correction of false results.

The adaption of the methods to other sensor data (f.e. RADAR data) and the exploitation of map data as an input to cartographic information systems is planned.

REFERENCES

BAUSCH, U., KESTNER, W.   1979 a, "Auswertung lokaler Zusammenhänge zur Bestimmung von
                          Objektgrenzen in Luftbildern", Informatik Fachberichte Nr. 20,
                          Springer Verlag Berlin

BAUSCH, U., et. al.,   1979 b, "Extraktion von Objekten aus Luftbildern durch objekt-
                          spezifische Verfahren mit stufenweiser Verbesserung der ört-
                          lichen Genauigkeit", Informatik Fachberichte Nr. 20,
                          Springer Verlag Berlin

BRACONNE, S., BRUN, F.,   1978, "Cartographie Automatique des Objects Lineaires",
                          Proc. Int. Conf. on Earth Observation from Space, Toulouse

DFVLR   1980, "Auswertung von Satellitenaufnahmen zur Gewinnung von
                          Flächennutzungsdaten für die räumliche Planung", DFVLR-IFP
                          Forschungsbericht, DFVLR Oberpfaffenhofen, in Press

GROCH, W.-D.   1979, "Zwei Verfahren zur vollautomatischen Suche von Start-
                          punkten für die Extraktion linienhafter Objekte aus Grauwert-
                          bildern", Informatik Fachberichte Nr. 20, Springer Verlag
                          Berlin

GROCH, W.-D.   1980, "Automatisierung der Extraktion linienhafter Objekte aus
                          Grauwertbildern"
                          Dissertation, University Karlsruhe, in Press

KENG, J.   1977, "Image Segmentation and Object Detection by a Syntactic
                          Method", Proc. Image Understanding Workshop, Minneapolis, USA

KYOTO   1978, 4. Joint Int. Conf. on Pattern Recognition, Kyoto, Japan

MONTOTO, L.   1977, "Digital Detection of Linear Features in Satellite
                          Imagery", Proc. Int. Symp. on Image Processing, Graz, Austria

NAGAO, M. et. al.   1978, "Region Extraction and Shape Analysis of Aerial Photo-
                          graphs", Proc. 4. J. Int. Conf. on Pattern Recognition, Kyoto,
                          Japan

NEVATIA, R.   1978, "Linear Feature Extraction", Proc. Image Understanding
                          Workshop, USA

OUAM, L.H.   1978, "Road Tracking and Anomaly Detection in Aerial Imagery",
                          Proc. Image Understanding Workshop, USA

OUIEL, F.   1979, "Luftbildinterpretation und multispektrale Klassifizie-
                          rung zur Gewinnung von Landnutzungsdaten", Bildmessung und
                          Luftbildwesen 47, Heft 4

ZUCKER, S.W., et. al.   1977, "An Application of Relaxation Labeling to Line and Curve
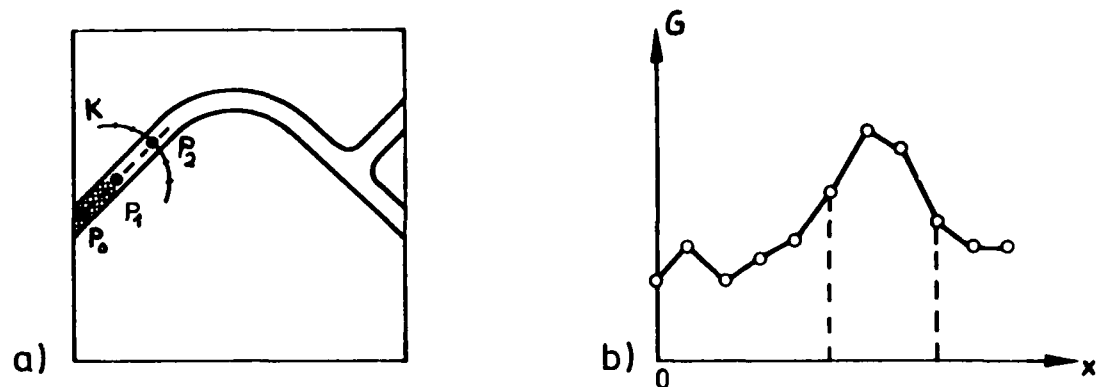   Enhancement", IEEE Trans. on Comp., Vol. 26, April 1977

Figure 1:   Basic proceeding of the incremental method;
             a) prediction of point P_2 ; b) grey level function of semi circle K

a)   b)   c)
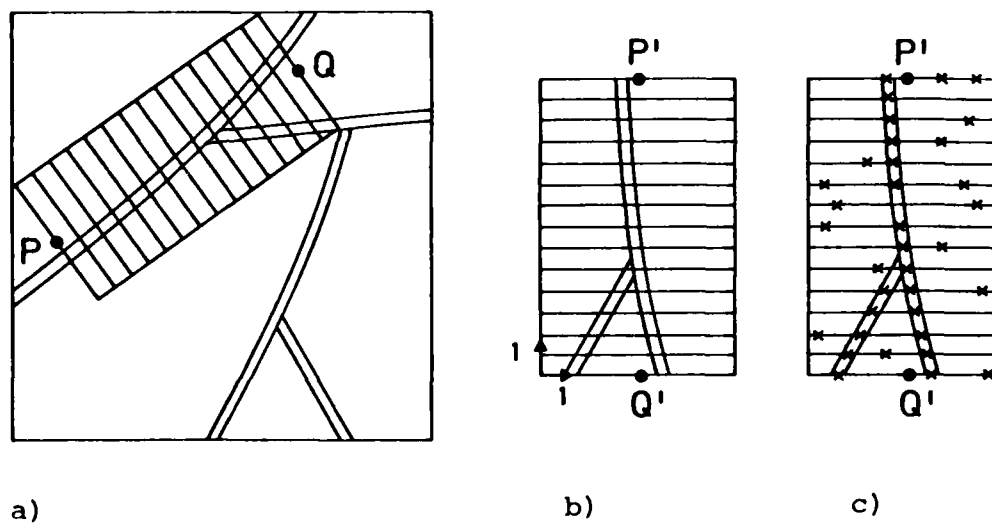
Figure 2:   Characteristics of the area slicing method;
             a) area of interest superimposed to the image; b) area of interest
             consisting of sample lines; c) locations with a high value of
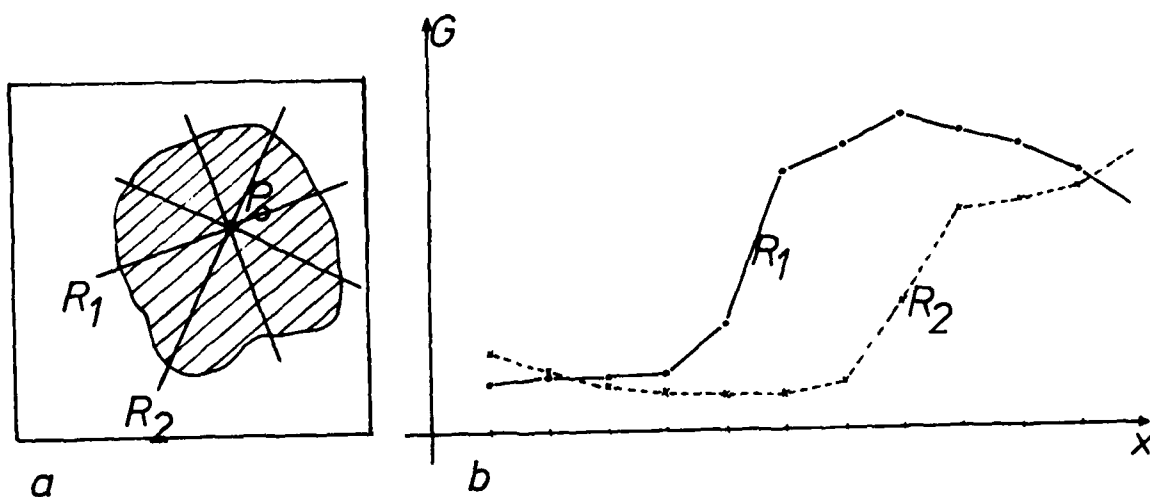   compliance C

Figure 3:   Basic idea of the radius method;
a) radii starting at central point P_0 ; b) grey level functions for
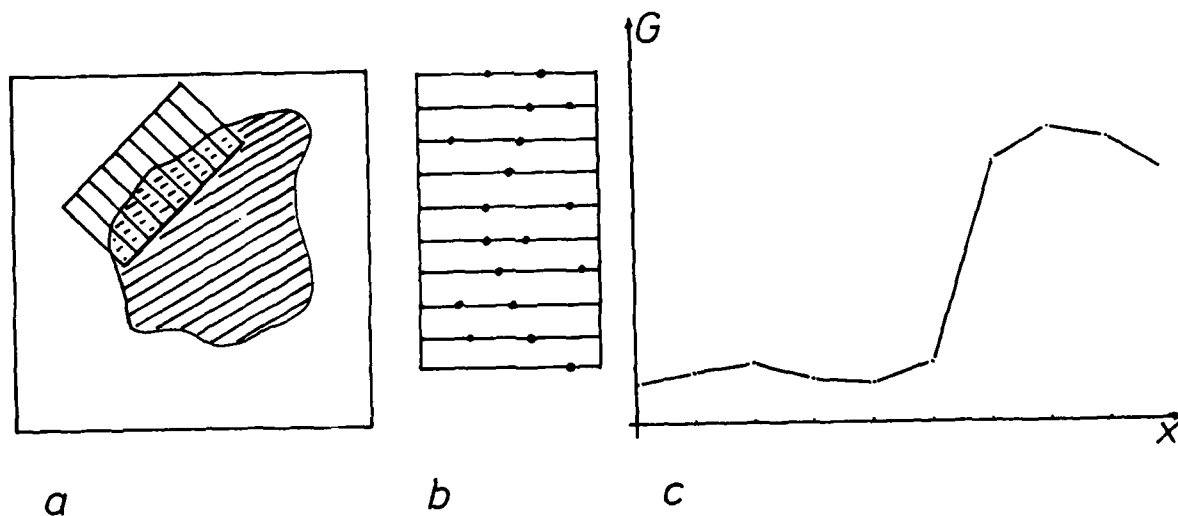radii R_1 and R_2

Figure 4:   The modified area slicing method;
a) area of interest containing a part of a region's contour; b) sample
lines with marked locations of high values C_m ; c) grey level function of
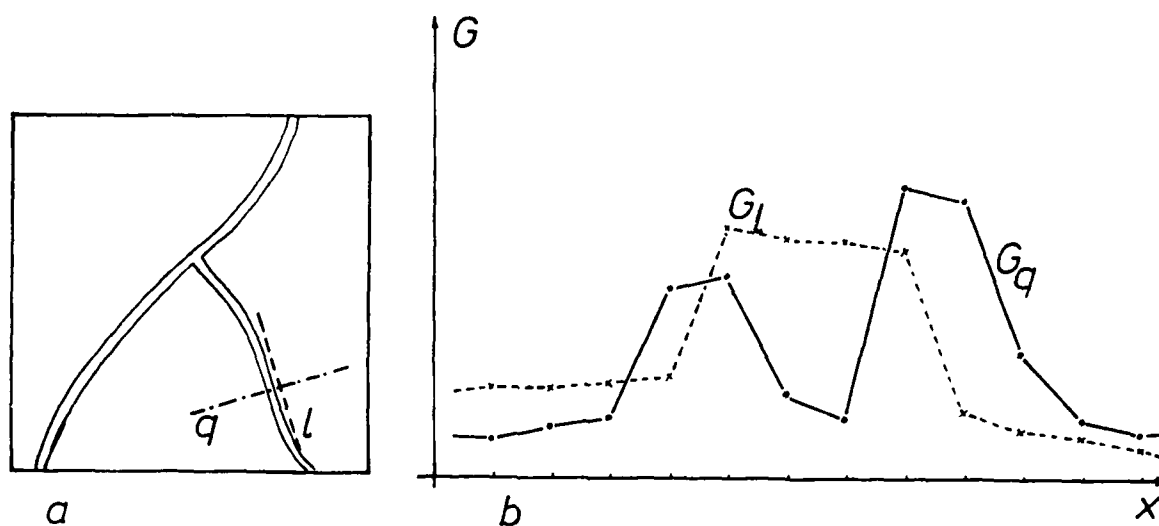first sample line

Figure 5:   To explain the extraction of point objects;
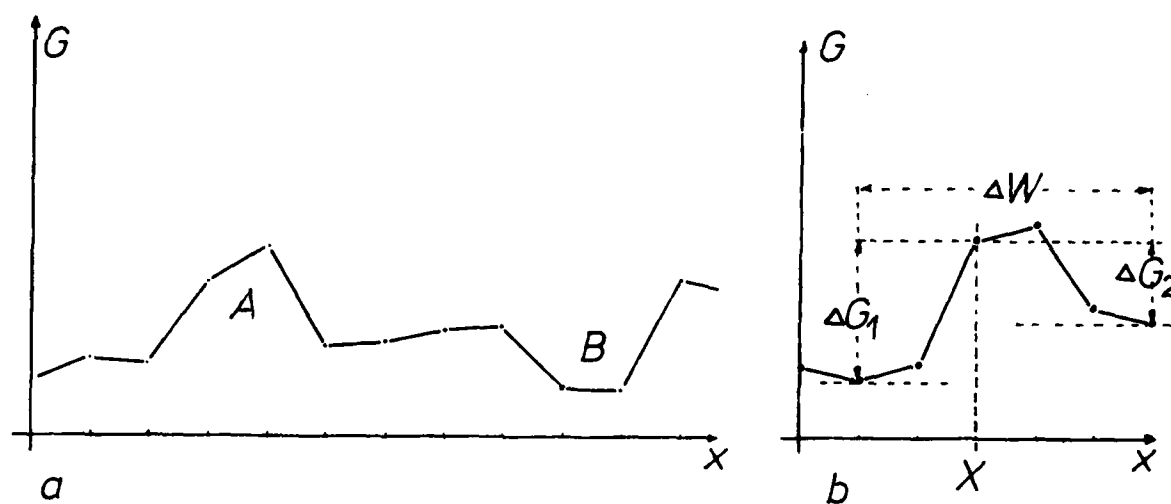a) test lines l and q; b) grey level functions of test lines

Figure 6:   Analysis of grey level profiles of line objects;
a) idealized profiles of line objects; b) features of the profile at
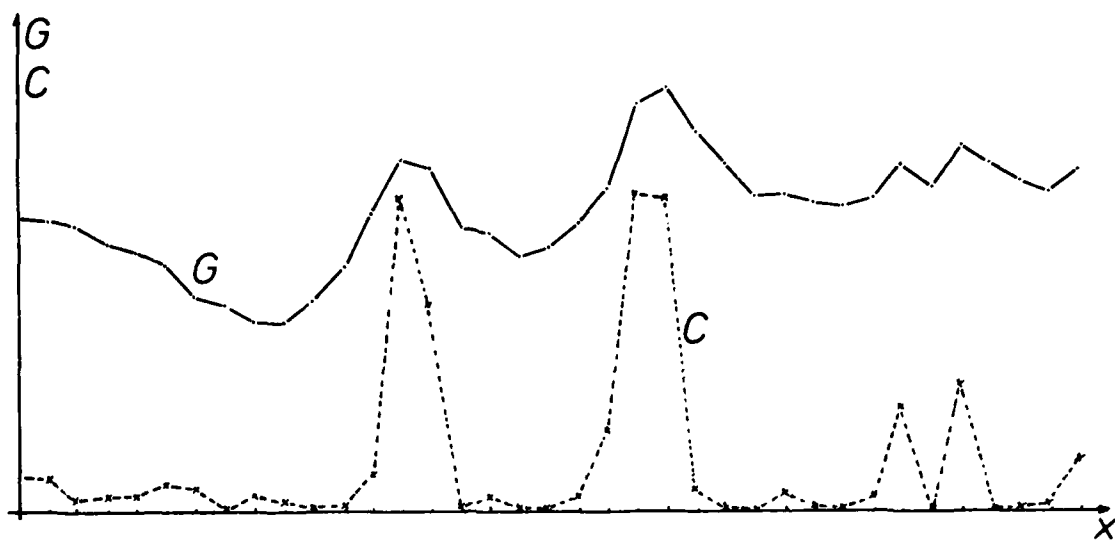location X

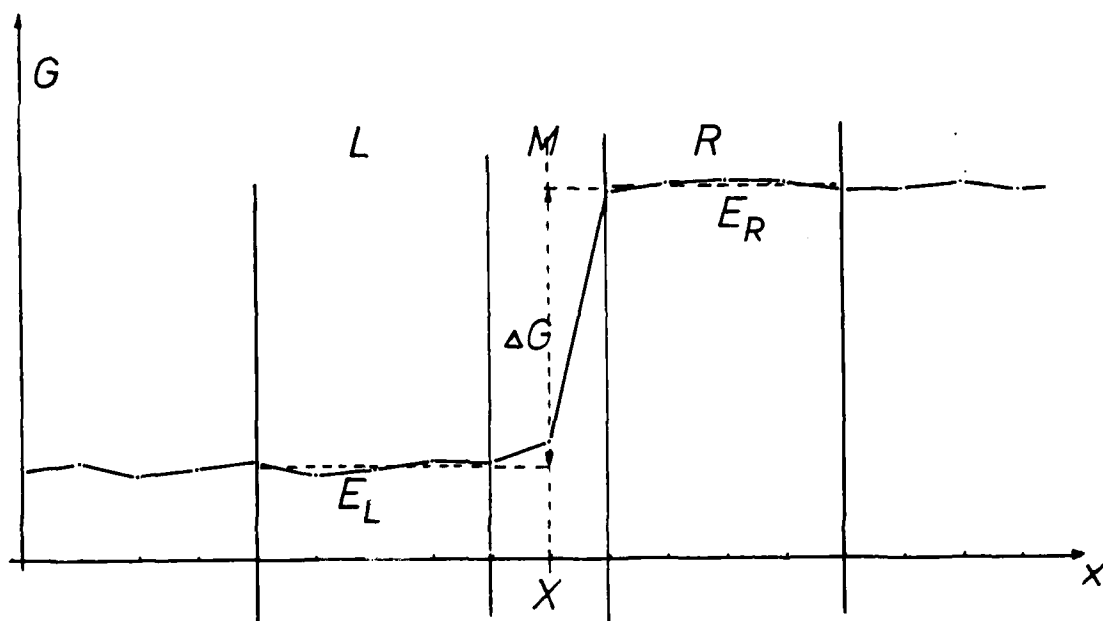Figure 7:   Demonstration of the line detection algorithm

Figure 8:   Analysis of grey level profiles of object contours

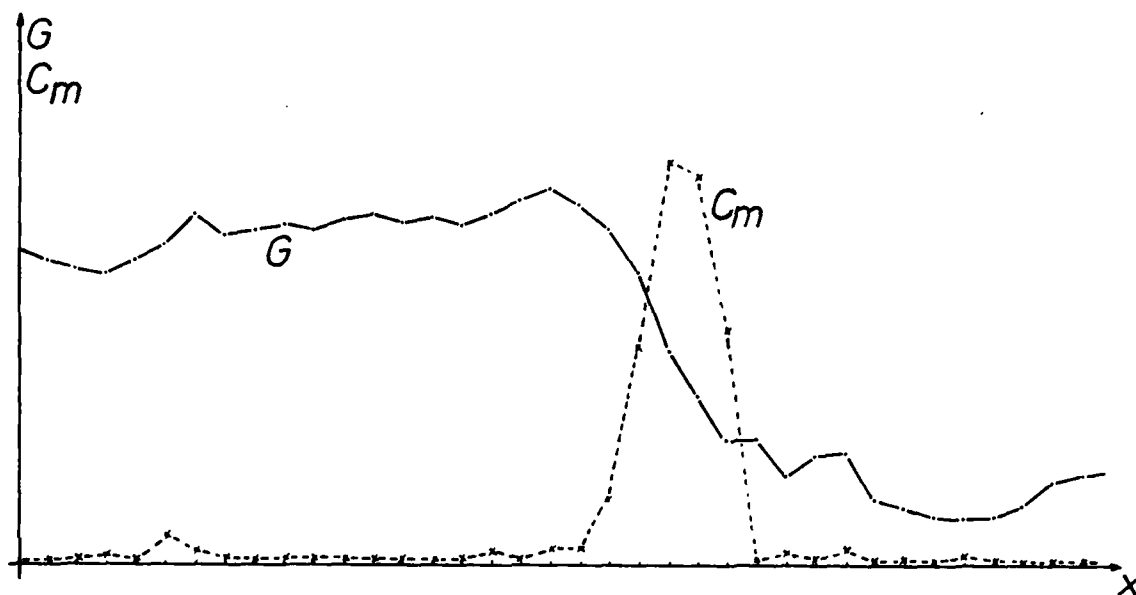Figure 9:   Demonstration of the contour detection algorithm
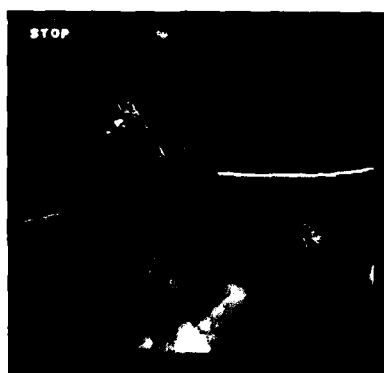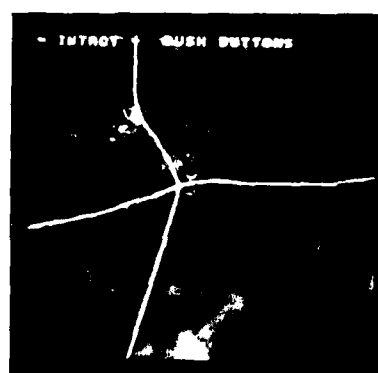
(a)   (b)

Figure 10:   Result of the extraction of roads by the incremental
             method

(a)   (b)

Figure 11:   Result of the extraction of various line objects

(a)   (b)

Figure 12:   Result of the extraction of a forest region

(a)

(b)

(c)

Figure 13: Result of the extraction of vehicles;
a) unprocessed parts of an image; b) cues found in the search areas;
c) accepted vehicle locations

A DETECTION AND PROCESSING SCHEME FOR MOVING TARGET DETECTION USING A PASSIVE SENSOR

Charles M Henderson
Royal Signals and Radar Establishment
St Andrews Rd
Malvern
Worcestershire
United Kingdom

SUMMARY

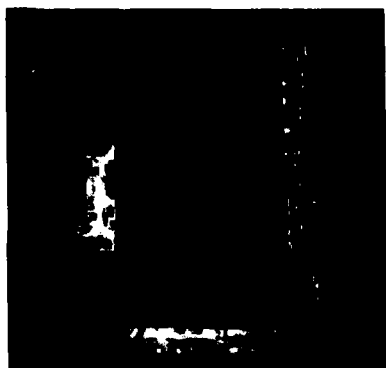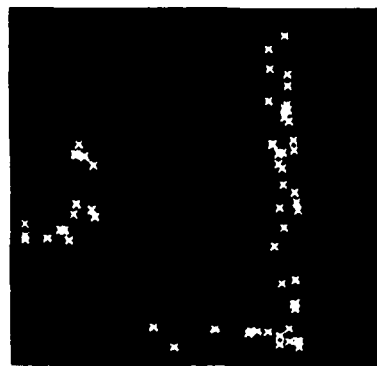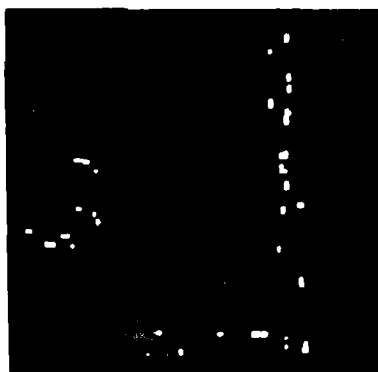An adaptive spatial filter has been implemented in digital hardware and a number of subsequent processing steps have been developed including static clutter cancellation and tracking. The hardware is part of an experimental ground based infra-red system for low level air surveillance.

The spatial filter algorithm is based on using the intensities recorded in a small window to detect locally significant peaks. The effect of noise, quantisation and mismatch between channels is discussed. Variation in target subtense, target registration and inter-channel gaps affect detection probability and this also has been explored by computer simulation.

The filter produces multiple detections on a single peak and therefore an additional processing stage was devised to reduce multiple detections to single detections. Stationary alarms are removed by a process of static clutter cancellation and remaining alarms are handled by track forming using a microprocessor. Angular rate limits are applied to tracks and established tracks, formed within rate limits, are confirmed as moving targets.

SYMBOLS

C	algorithm 1 threshold parameter
K	algorithm 2 threshold parameter
x_n	co-ordinates of an alarm on the nth scan
px_n	predicted position of an alarm on the nth scan
Δx	prediction window half width
CE	central pixel recorded amplitude
V_i	pixel element i recorded amplitude

1.   INTRODUCTION

This work has been carried out in the context of the feasibility of passive, automatic infra-red surveillance for short range surface to air guided weapons. The emphasis of the work is towards achieving around 10 km detection range against fighter ground attack aircraft at a operating false alarm rate of 2-3 per hour.

Initial work was aimed at providing a raw data base for targets and backgrounds but the scope of work was subsequently widened to provide an experimental hardware implementation. An experimental panoramic surveillance equipment operating in the 8-13μm waveband was procured from EMIE at Feltham. Concurrent with the sensor procurement Plessey at Havant were contracted to investigate signal processing schemes, in particular digital spatial filter algorithms, and to produce a hardware implementation of the most promising spatial filter algorithm studied. The filter hardware was interfaced to the EMIE scanner and the RSG9 trial at Taranto 1978 was the first opportunity to demonstrate the techniques evolved and to gather data suitable for further investigation. As a result of further work additional processing hardware has been built to complete the system as described.

1.1   The Sensor

Fig 1 shows a functional block diagram of the equipment. The sensor provides image quality data over a coverage of 360° x 6.8° at a revolution rate of twice per second. The elevation coverage of the equipment is achieved in a single swathe per scan using a 192 element detector at a geometrical resolution of 0.5mR. The figure shows that two outputs of the head are used. The whole panoramic coverage is digitised to 8 bit resolution and the resulting real time digital video made available to the signal processing. Secondly, a sector of the panorama covering 13.4° x 6.8° is stored to 6 bit resolution for TV display. The bearing of the sector being selectable over 360°.

The sensor head achieves panoramic coverage by rotating a 45° mirror in a periscope structure. The detector with its associated cooling system is fixed and a stationary image is presented to the image plane by means of a sophisticated derotation system. The detector output signals are AC coupled to pre-amplifiers and after band shaping,amplification and multiplexing are converted to 8 bit samples for signal processing. Each channel is sampled at 0.49 milli radian intervals in azimuth giving a data rate of 4.9 x 10^6 pixels per second.

1.2   The Spatial Filter

With a 8-13μm broadband surveillance sensor, spatial filtering provides the prime technique whereby an infra-red target can be distinguished from the majority of clutter background detail. The statistics of IR background signals are strongly non-stationary and attempts to design a single optimum filter are doomed to failure since the radiance distributions change markedly with elevation (ground, horizon, sky) azimuth angle (buildings clustered over small sectors) and time (solar heating, deployment of battlefield

paraphernalia). What is needed is an adaptive filter capable of adjusting its threshold decision as the need arises. The spatial filter algorithm used is based on using the intensities recorded in a small window to detect locally significant peaks. The window is a 3 x 3 pixel grid, which at its most compact is a sparse arrangement within a 3 x 5 rectangle. The detection criteria are that the central pixel amplitude shall exceed the mean of the other grid elements by a preset value C and exceed the mean plus a preset fraction K of the difference between the maximum and minimum element values. These thresholds are calculated and refreshed at each and every pixel during the surveillance scan. In this way, the decision threshold instantaneously applied to each pixel in turn varies with the clutter content in the neighbourhood of the pixel and by choosing both the overall size of the window and its element sizes, small hotspots including aircraft targets are extracted. Relatively wide angle detail (cloud edges, masts etc) or adjacent conglomerations of hotspots are rejected. The output of the spatial filter consists essentially of the co-ordinates of detected hotspots, but other data used in the filter including amplitude is also available.

Being a passive broadband sensor this information is not supplemented by either range or velocity data.

1.3   Subsequent Processing Stages

Recorded output from the spatial filter formed a data base for computer simulation of subsequent processing stages which have now been implemented, and figures quoted for data reduction by subsequent stages are those obtained from simulations on an extensive data base. It was originally conceived that the processing of alarms should be handled by a single microprocessor producing tracks and flagging those tracks which were target like. This leads to the need for a degree of algorithmic sophistication which meant that the processor used was not capable of handling the data rate. By adding intermediate processing stages it was possible to reduce the load on the processor and simplify the tracking algorithms to allow for an increased data rate and possibly some enhancements to the basic tracking concept.

2.   ANALYSIS AND RESULTS

2.1   Spatial Filter

The general configuration of the spatial filter is shown in Fig 2. The algorithm was chosen because of the simplicity with which it can be implemented digitally. In the simple case where signal energy is concentrated in the central element and the surrounding 8 pixels are each independent samples of the same gaussian noise process the probability of detection and false alarm on noise is analytic (Appendix I). The two thresholds calculated by the filter are distributed with mean values of C and 2.845K referred to rms noise. The distribution of the second threshold is skew with long tails which means both a high probability of false alarm on noise and a high probability of setting a high threshold on signal, neither of which is particularly desirable. The distribution of the first threshold is a narrow gaussian and was introduced as a 'catch' to reduce the false alarm rate on noise (fig 3) which results from the other threshold. The practical situation does not lend itself so readily to analysis. The 8 pixels used in the algorithm to derive a threshold which signal must exceed are taken from 3 separate detectors. Although these are approximately matched by the manufacturer, there is none the less a spread in their detectivities and responsivities and a spread in the matching of the gains and offsets of their respective preamplifiers caused by misalignment, ageing and thermal drift. The noise processes which occur within the dewar containing the detectors combine in such a way that the assumption of independence between samples is questionable. The filter hardware uses discrete digitised samples to perform its arithmetic and hence the dynamic range of signals is limited. A Monte Carlo model approach was adopted in an attempt to assess the probable effect of mismatches and to see at what level of quantisation the discrete solution differed from the continuous analytic solution. For quantisation levels less than rms noise there is little difference between the discrete and continuous performance but beyond that the K threshold ceases to be significant and hence the ability of the threshold to effectively adapt to background detail. Taking account of this and the maximum quantisation error in a signal it is undesirable to quantise more coarsely than around 1 level for rms noise. The effects of the degradation processes in the real situation are an increase in false alarm rate and a reduction in target detection probability as one might expect. See Fig 4. There is an increasing probability that a large signal will not be detected, as the system is progressively degraded by mismatching, this relates physically to the banding which is seen with parallel scan thermal images and may, to some extent, be compensated for by varying the parameters K and C dynamically. This has been allowed for in the implementation and K and C may be altered at video rates.

The spatial filter in the absence of noise is capable of discriminating against those features which would give rise to an alarm with a simple one dimensional filter, see fig 5, and in favour of point targets. The subtense of a true target however depends on its range, size and aspect and is only unresolved at the furthest range at which the sensor is capable of detecting it. Figure 6 illustrates the spatial properties of just two possible windows against a square target. The detector output signals are calculated by convolving the target with the optics point spread function and the detector, and the signals presented to the filter are calculated by convolving the detector output signals with the impulse response of the electronics. To obtain the probability of detection the square target is taken to be centred with equal probability anywhere within a rectangle, centred on the central element of the window, 0.49 mR wide by 0.625 mR high, to account for registration losses. The size of the rectangle is fixed by the sampling rate and the detector pitch. As target subtense increases it spreads into the 8 pixels surrounding the central element increasing their mean value and hence raising the detection threshold. Because the target may not register perfectly with the sampling and scanning the spread is preferential to one or other element, this also raises the detection threshold depending on the value of K. The first window is not a suitable choice, K must be near one at least to discriminate against simple features and that value leads to poor target detection performance. Window 2, however, has good performance for targets less than 1.5 mR subtense and discriminates well against targets of larger subtense. In the case of small signals and in the presence of noise both thresholds must be considered as the target size is increased. For small targets misregistration between target and signal sample leads to a reduced central element value, and spreading of larger targets leads to an increase in the mean. These two mechanisms

increase the effective value of C, see Fig 7. The effect of large target subtense in K has been accounted for in Fig 6 but the effect of small target loss has not, it is in fact simply a reduction in signal to noise ratio.

2.2   Multiple Alarm Cancellation

Because the spatial filter uses a sparse matrix of picture elements but computes for every picture element there is a possibility that a point or slightly extended source will generate more than one alarm, the number of alarms possible from a single source increases as the window size and as the source size increases within the limits of the window size. In fact for window 2 around 30% of all clutter alarms consist of clusters of alarms. Aircraft target alarms are seemingly more likely to be multiples. From an analysis of the RSG9 Taranto trial results 60% of aircraft detections at 10 km range consisted of more than one alarm and 95% at 5 km consisted of more than one alarm.

These alarms have more than a nuisance value in increasing the loading on the tracking process, they lead to the creation of more than one trackfile for what is essentially the same track. This leads in turn to conflicts in associating new plots with existing tracks and an unacceptable increase in the target track confirmation time. There is no penalty in reducing connected clusters of alarms to single alarms.

Two algorithms have been studied. The first is based on rejecting all but the alarm with the largest amplitude of a set of connected alarms. Alarms are reckoned to be connected if their co-ordinates are within a predetermined range of each other, say 2.5 milli radians. The advantage of this algorithm with targets is that it tends to determine the thermal centroid and it may reduce the number of clutter false alarms generated by objects in the foreground such as wire stays which a sparse filter matrix does not effectively reject. The disadvantage of the algorithm is that it is not easily implemented. The second algorithm is to take the first alarm detected and inhibit further alarms within a rectangular window of which the first alarm is the corner. This algorithm is easily implemented.

A comparison showed that the first algorithm was only marginally more effective than the latter and that a window of 5 x 5 pixels in the second multiple cancelling algorithm is near to optimum. This has now been translated into hardware.

2.3   Static Clutter Cancellation

Having removed multiple alarms from the output of the spatial filter there are inevitably other alarms originating from detail in the thermal scene which are essentially stationary. If trackfiles are set up to track these stationary targets then an unnecessary load is placed on the processor. They do in fact warrant a separate process which has been implemented after multiple alarm cancellation. The mechanism involves the storage of a map of alarms and correlating incoming alarms with the map and at the same time updating the map. The reason for implementing this process after multiple alarm cancellation is that it reduces the amount of storage needed for the map. Correlation is carried out to a lower precision than the angular data rate, this is to take care of scan to scan angular jitter from sensor registration noise, wind blown clutter and apparent movement of highlights on a geometrically extended source.

The static clutter cancellation works as follows. The coordinates of all the alarms received on the first scan are stored sequentially. A number is also stored with each alarm denoting the number of scans over which that alarm has been seen, this is called the clutter number. In the first instance this would be one. On the second and subsequent scans alarms received are compared with the stored alarms and new alarms added to the store, alarms already in the store from previous scans have their clutter number incremented by one, up to a predetermined limit and stored alarms which have not been sensed have their clutter number decremented. Predetermined upper and lower limits decide whether the alarm is passed on to the tracking computer or deleted from the clutter store.

By selecting the limits and increments it is possible to adjust the number of times that an alarm must be seen before it is inhibited from the next processing stage of tracking and the number of times that a previous static alarm may fail to be seen before it is deleted from the store. The reason for these criteria are so that targets which are slow moving are not masked nor delayed from being confirmed and that scintillating clutter is not continually being deleted and re-added to the map.

The efficiency of such a scheme will depend strongly on the scene being surveyed but on simultations using data gathered at RSG9 Taranto trials it effected between a 50% and 70% reduction in the number of alarms. The choice of resolution and the number of times an alarm must be seen before it is inhibited is dependent on the slowest targets which it is desirable to track, around 1 milli radian per second. A window size of 1 mR for correlation around each alarm was chosen with the requirement that an alarm be seen 3 times before it is cancelled. The choice of parameters is not critical, halving the resolution yields only a 2% improvement in cancellation rate and increasing the time to delete an alarm from store from 2.5 seconds to 25 seconds yields only a 1% improvement. The software implementation of the algorithm uses a linked data structure with pointers to enable rapid and efficient correlation, the translation into hardware has been made using two stores and associated logic, each of 1000 alarm capacity which alternate in use every scan. One is the reference map which is read from and the other the updated map which is written into from the old map and newly received alarms. All new alarms and those not yet inhibited are passed to the microprocessor for track forming.

2.4   Microprocessor Track Forming

Remaining alarms are processed by a microprocessor to form tracks. The idea is to associate the angular pattern of alarms from scan to scan and deliberately search for movement by establishing tracks from scan to scan plot data. As each track is updated, confidence in deciding that the ttack belongs to a target is updated according to preset algorithms and at some point in the processing the presence of a target is confirmed.

The basic approach is to establish a trackfile for each new, unassociated alarm or plot found in each scan. For example at the end of a scan there may be say five trackfiles, on the next scan there may be ten plots present. The task is to associate the existing trackfiles with the plots present. Association is carried out by storing in each trackfile a predicted track position for the next scan together with an association window. If one of the next scan plots falls within the prediction window of a trackfile it will be associated with the track and the trackfile updated. The way in which the trackfile is updated depends on the results of the association process. Eventually, in the case of a consistent target track, sufficient confidence is accumulated to cause track confirmation or trackfiles are deleted and never confirm because of lack of updating and lack of confidence. As the trackfiles become regularly updated, targets can be classified as fast or slow and may therefore be deleted on the assumption that they are new fixed clutter which has not been inhibited by the static clutter cancellation hardware. The actual association is attempted by creating a shortlist of plots which fall within the prediction window of each trackfile. If no plots are present the window is doubled in size and the process repeated. If there are still no associations found, the track is coasted and the confidence factor decremented. If the window size has not been widened and a single plot is found in the shortlist, this represents the optimum situation and the confidence factor is incremented. Table 1 summarises the rules used to cover all eventualities during the association process. Following association the trackfile is updated by re-filing the track history over the last 3 scans and re-computing a new prediction for the next scan. The prediction rules are:

one point prediction (new trackfile)

$$px_{n+1} = x_n \pm \Delta x$$

two point prediction

$$px_{n+1} = 2x_n - x_{n-1} \pm \Delta x/_2$$

three point prediction

1) track not coasted

$$px_{n+1} = 3x_n - 3x_{n-1} + x_{n-2} \pm \Delta x/_6$$

2) track coasted

$$px_{n+1} = 3x_n - 3x_{n-1} + x_{n-2} \pm \Delta x/_3$$

where px_{n+1} is the predicted position of the next alarm of an established track, Δx the half width of the association window and x_n is the position on the nth scan. It can be seen that the prediction window size is adapted approximately as more information on the track becomes available. The initial value Δx (1 point prediction) is set so that targets with radial rates greater than 3.5 degrees per second are not tracked.

When all trackfiles have been updated for each alarm not used in the association and updating process a new trackfile is established. Finally all the trackfiles are examined for status to determine whether the track may be confirmed or deleted. For a track to be confirmed as an alarm it must have reached a predetermined confidence level, the slow target flag must not be set and the coast flag must be cleared. Tracks are deleted if they have to be coasted for two scans or if the confidence falls to zero. In general for a well behaved track confirmation will occur after three consecutive detections within four scans. This is essential to maintain reasonable target confirmation ranges. A slow target is not deleted so long as it continues to be updated but unless its status changes it is never confirmed.

3. CONCLUSIONS

On the basis of measured performance of the spatial filter it has been possible to classify false alarms into distinct types and postulate means for eliminating most of them. The key process is the spatial filter itself and its ability to select target structures. The massive data reduction that spatial filtering produces makes the problem of scan to scan correlation more tractable and this in turn reduces the data rate to manageable properties for kinematic track processing. Indeed each step in the pipeline is designed to reduce the load on the next more sophisticated step. In this way the task has been realised in real time without resorting to high technology or very high cost structures.

APPENDIX 1. CALCULATION OF SPATIAL FILTER PERFORMANCE

The spatial filter uses a matrix of 9 pixels, a central element CE and 8 neighbours $V_1 \ldots V_8$. Assuming that each pixel value is an independent sample of gaussian noise with unit variance and zero mean and $V_1 \ldots V_8$ are numbered in order of increasing amplitude the two algorithms used in the filter are.

$$CE > \frac{1}{8} \sum_1^8 V_i + C \tag{1}$$

$$CE > \frac{1}{8} \sum_1^8 V_i + K(V_8 - V_1) \tag{2}$$

The frequency density function of V is

$$f(V) = \frac{1}{\sqrt{2\pi}} e^{-\frac{V^2}{2}}$$

and the probability of V_i having the value u in the range u to u + du on a single sample is f(u)du. The mean value of a set of samples of gaussian noise is itself normally distributed about the population mean with a frequency density function $f_m(V)$, which for 8 samples is

$$f_m(V) = \frac{2}{\sqrt{\pi}} e^{-4V^2} \tag{3}$$

The pixel CE also contains a noise sample which may be shifted to the right of equations (1) and (2) and subtracted from the mean.

The frequency density function f_1 of the thus modified mean is the convolution

$$f_1(V) = \int f_m(\phi) * f(\phi - V) \, d\phi \tag{4}$$

This evaluates to a gaussian distribution

$$f_1(V) = \frac{1}{\sqrt{2.25\pi}} e^{-\frac{V^2}{2.25}} \tag{5}$$

Given a signal of amplitude V_T in the central element the probability of detecting it with the first algorithm is the probability that $\hat{V}_T > V_m' + C$ which is

$$Pd1'(V_T) = \int_{-\infty}^{VT+C} f_1(x) \, dx \tag{6}$$

and the probability of a false alarm on noise in a single sample

$$Pfal = \int_{-\infty}^{-C} f_1(x) \, dx \tag{7}$$

The frequency density function of $V_8 - V_1$ may be derived as follows. The probability of taking a sample in the range u – u + du is f(u)du and in the range w – w + dw is f(w)dw. The probability that a sample will be taken between u and w

$$g = \int_{u+du}^{w} f(x) \, dx \tag{8}$$

Consequently the probability that, in selecting N samples, one will lie in the interval du, one in the interval dw and the remaining N-2 in the intervening interval is

$$dP = N(N - 1) \ f(u)f(W)g^{N-2} \, dudw \tag{9}$$

since N(N – 1) is the number of ways in which one of the values may fall in each of the intervals du and dw. Rearranging by making v = w – u and summing over all intervals du the frequency density function f_{RN} of Vn – V_1 is

$$f_{RN}(r) = N(N - 1) \int f(u)f(u + r) \ (g(u + R) - g(u))^{N-2} \, du \tag{10}$$

This is a skew distribution function with a mean value for N = 8 of 2.845.

The frequency density function of thresholds in algorithm 2, f2, may be derived by convolving (5) and (10)

$$f_2(V) = \int f_1(\phi) * f_{R8}\left(\frac{V - \phi}{K}\right) \, d\phi \tag{11}$$

This integral must be solved numerically.

The probability of detecting a signal V_T with this algorithm is the probability that $V_T > Vm' + K(V_8 - V_1)$ which is

$$Pd2(V_T) = \int_{-x}^{V_T} f_2(x)\ dx \qquad (12)$$

and the probability of false alarm a noise is

$$Pfa2 = \int_{-x}^{0} f_2(x)\ dx \qquad (13)$$

For an alarm to be registered both of algorithms (1) and (2) must be satisfied. A suitable choice of K for spatial discrimination is between 1 and 1.5, C is chosen to set the noise false alarm rate, see Figure 3.

TABLE 1

Track Association Rules

ACTION TAKEN FOR

PLOTS IN SHORTLIST	ORIGINAL PREDICTION WINDOW SIZE	DOUBLED WINDOW SIZE
0	DOUBLE WINDOW AND TRY AGAIN	COAST TRACK CONFIDENCE -5
1	CONFIDENCE +10	CONFIDENCE +5
2	TAKE NEAREST TO PREDICTED COORDINATE CONFIDENCE +5	TAKE NEAREST CONFIDENCE +5
3	TAKE NEAREST CONFIDENCE +5	TAKE NEAREST CONFIDENCE +5
> 3	CONFIDENCE -10	CONFIDENCE +5

PREDICTION WINDOW ΔX = 30 PIXELS

SLOW TARGET RATE LESS THAN 1 PIXEL MOVEMENT
PER SCAN (2 PIXEL IN AZIMUTH)

CONFIRMATION REQUIREMENTS
CONFIDENCE = 25
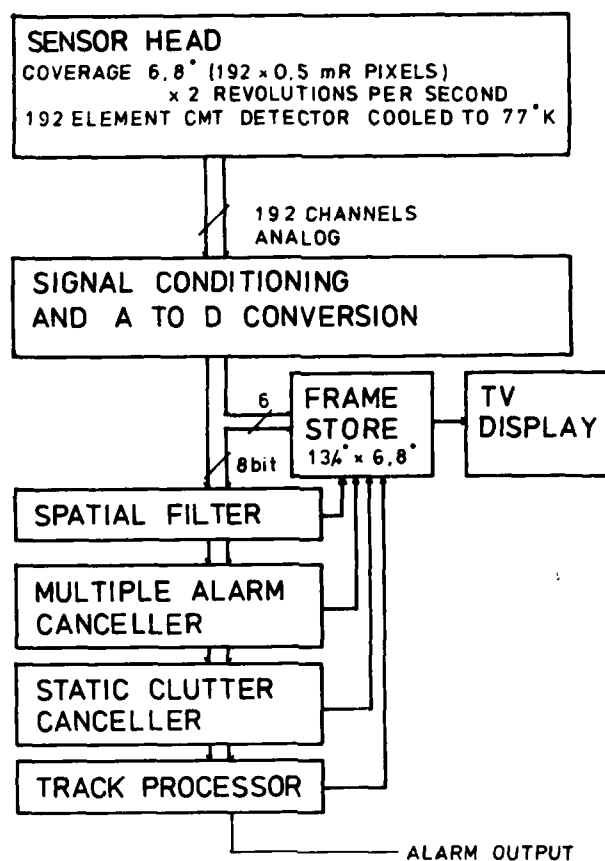COAST FLAG UNSET
SLOW TARGET FLAG UNSET

SENSOR HEAD
COVERAGE 6.8° (192 × 0.5 mR PIXELS)
x 2 REVOLUTIONS PER SECOND
192 ELEMENT CMT DETECTOR COOLED TO 77°K

192 CHANNELS
ANALOG

SIGNAL CONDITIONING
AND A TO D CONVERSION

6

FRAME
STORE
134° × 6.8°

TV
DISPLAY

8 bit

SPATIAL FILTER

MULTIPLE ALARM
CANCELLER

STATIC CLUTTER
CANCELLER

TRACK PROCESSOR

ALARM OUTPUT

Fig. 1   Block diagram of passive experimental
infra-red surveillance system

0.49 mrad

centre pixel
moves sample
by sample

DETECTOR

CHANNELS

0.625 mrad

WINDOW 1

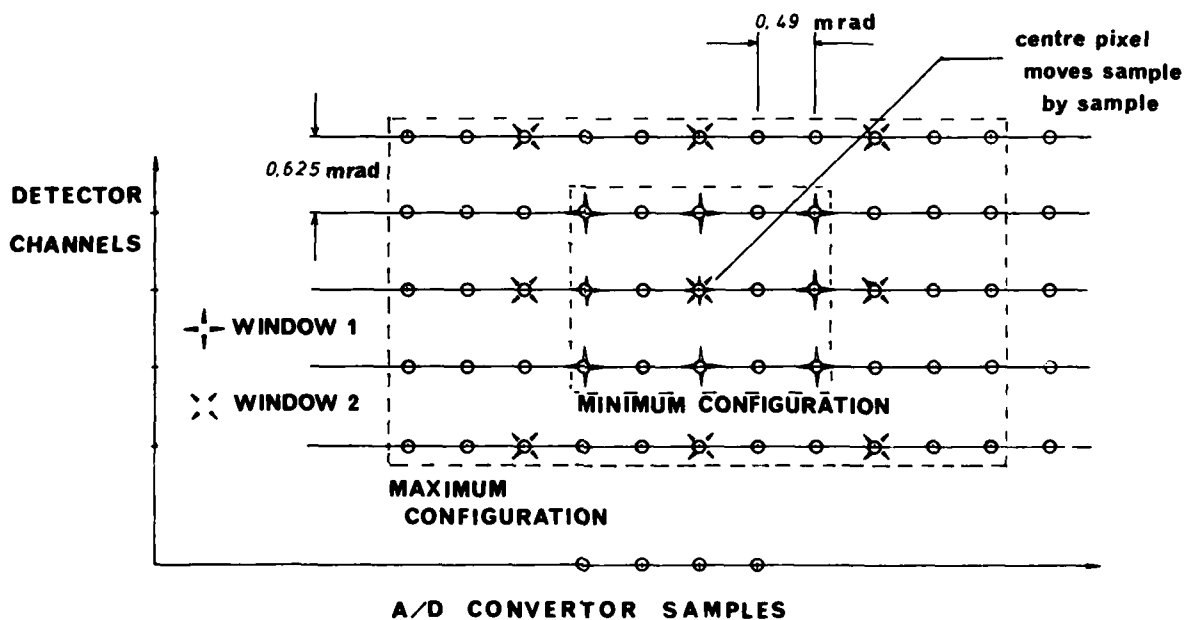WINDOW 2

MINIMUM CONFIGURATION

MAXIMUM
CONFIGURATION

A/D CONVERTOR SAMPLES

Fig.2   Spatial filter general configuration

probability
of false
alarm per
pixel

10^{-3}

10^{-4}

10^{-5}

10^{-6}

PfanC

PfanK

1   2   3   4   5   6   7
K . C x RMS Noise
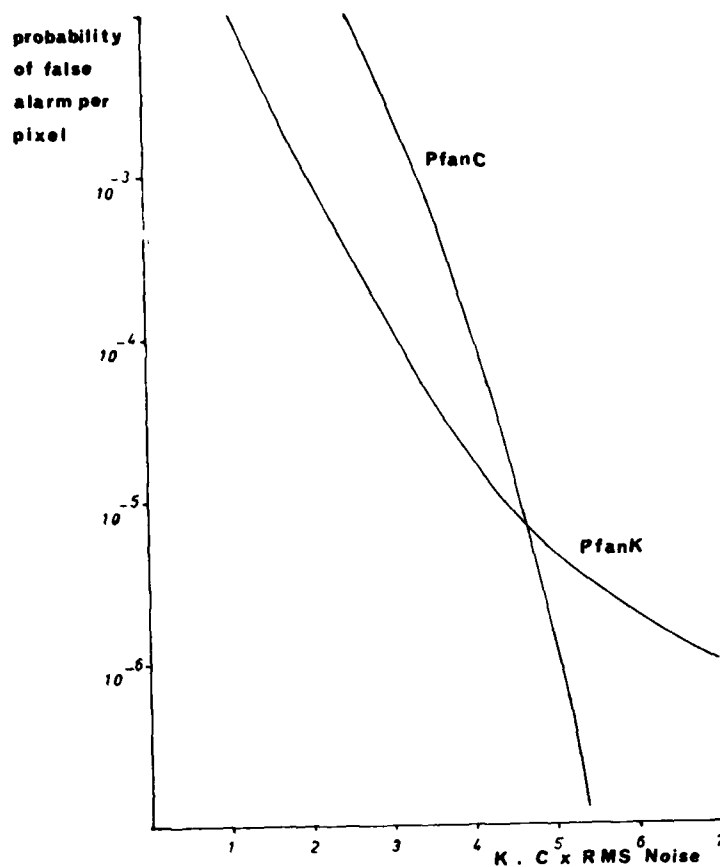
Fig.3   Probability of false alarm on noise vs K and C

PROBABILITY
OF
DETECTION

1

a IDEAL
K=1
C=6

b NON
IDEAL
K=1
C=6

c COARSE
QUANTIZED 4σ
K=1
C=6

d IDEAL
K=1,5
C=6

e NON IDEAL
K=1,5
C=6

.5

0

0   5   10   15   20
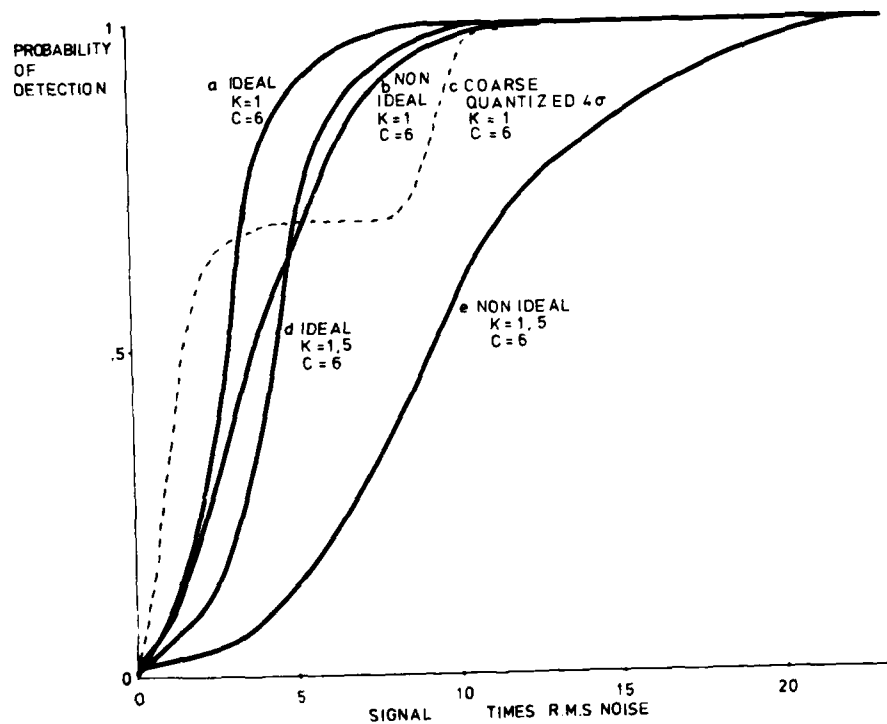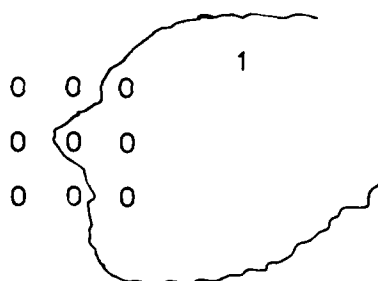SIGNAL   TIMES R.M.S NOISE

Fig.4   Detection probability against signal to noise ratio showing the degradation in performance
caused by over coarse quantisation and defects expected of a real system

THE ALGORITHM :   CENTRE ELEMENT >   MEAN + K(MAXIMUM
- MINIMUM)
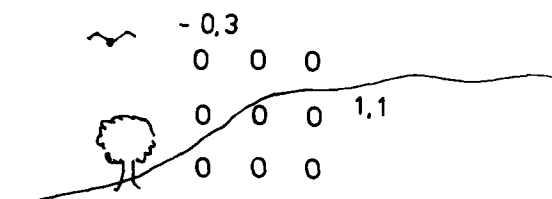PARAMETER K IS USUALLY SET IN THE RANGE 1 - 2

1 eg cloud edge

CE = 1   MEAN = 3/8
MAX =1   MIN   = 0

K > 0,625

NO ALARM

2 eg horizon
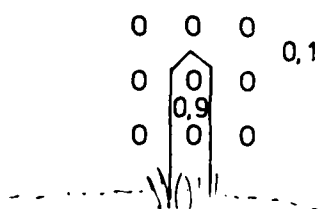

CE = 1,1 MEAN = 0,7
MAX =1,1 MIN   = -0,3

K > 0,285

NO ALARM

3 eg fence post top

CE = 0,9 MEAN = 0,2
MAX = 0,9 MIN = 0,1

K > 0,875

NO ALARM

Fig.5   Examples of the 3 x 3 window spatial filter
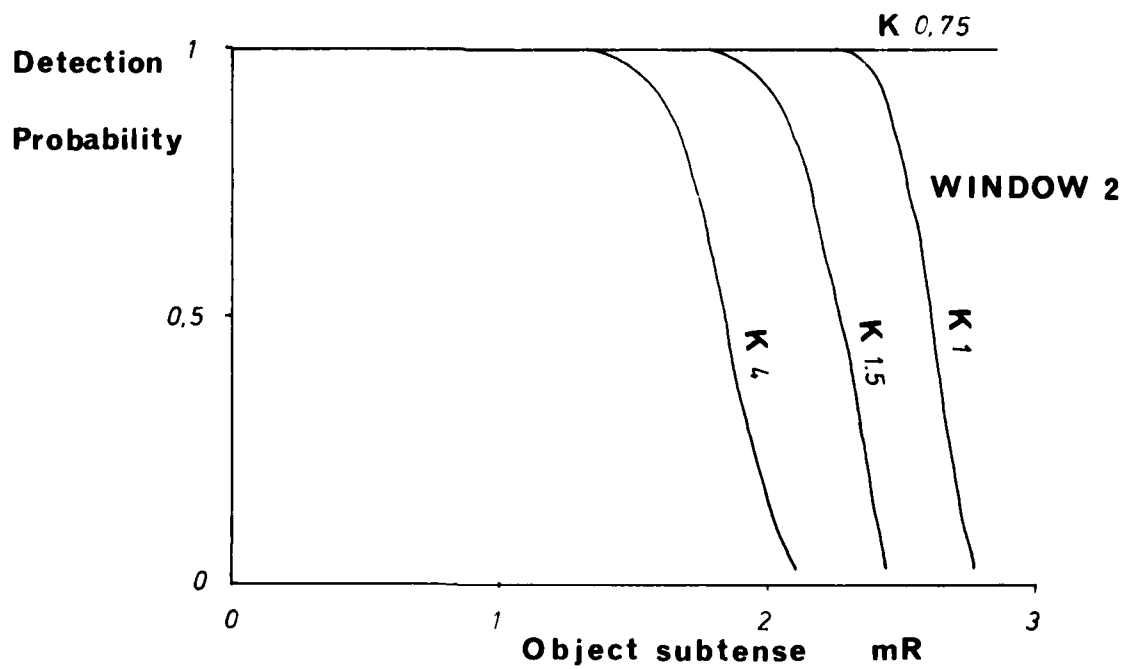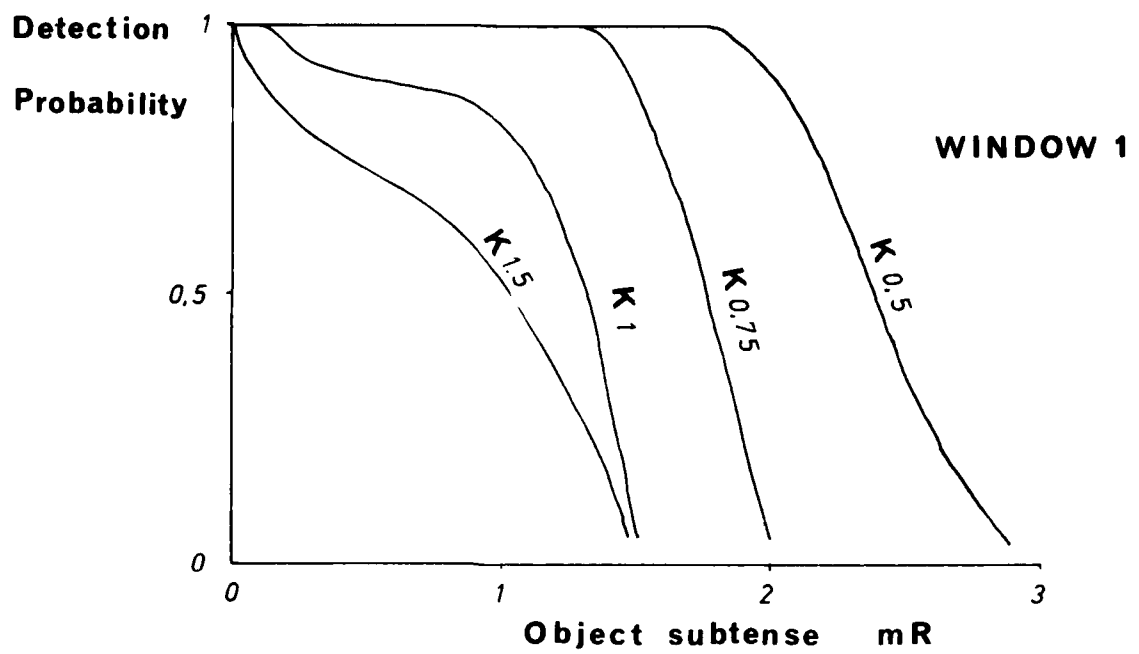and its ability to reject features

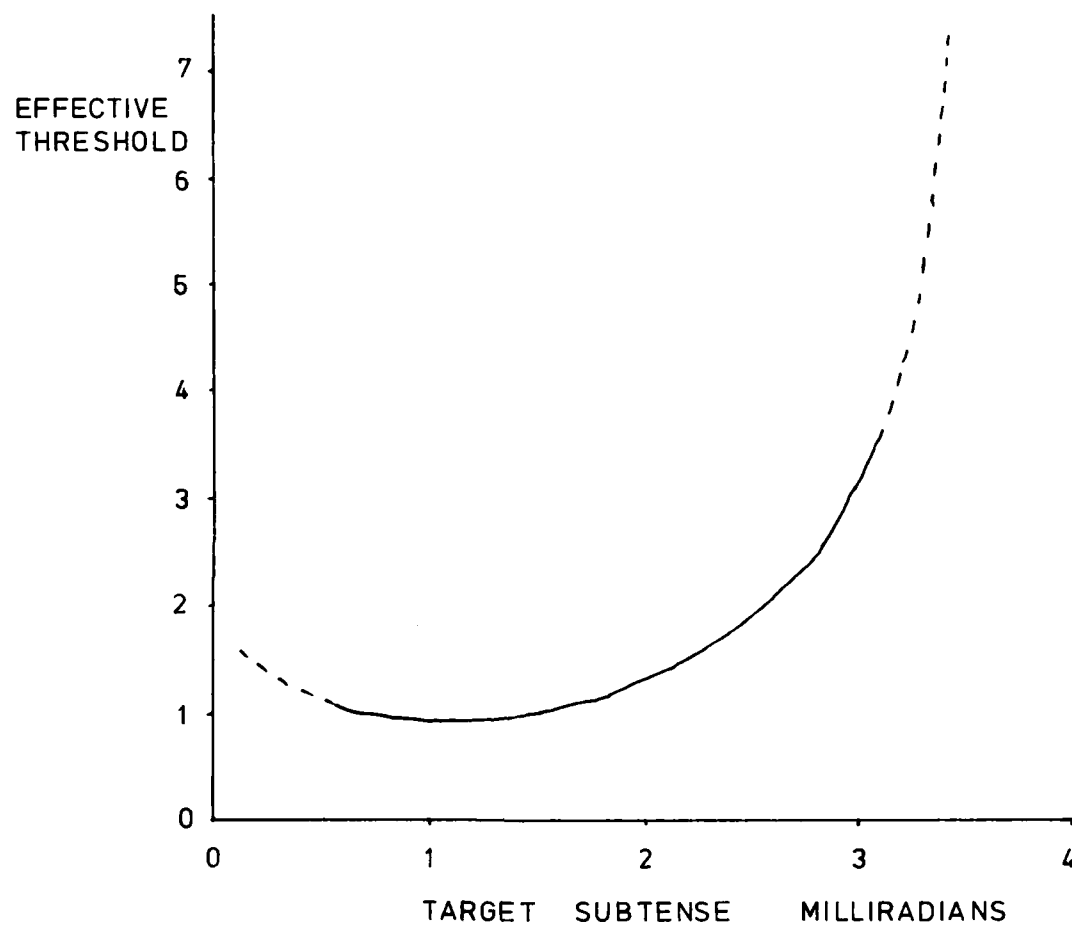Fig.6   Variation in filter performance with target size

Fig.7   Variation in effective (non-adaptive) threshold C with target subtense window 2
based on 80% acceptance by adaptive threshold K

A TV-Tracking System Based on Computer Intelligence

Dr. Gerhard Hirzinger, Ing. (grad.) Klaus Landzettel

Deutsche Forschungs- und Versuchsanstalt für Luft und Raumfahrt e.V. (DFVLR)
Institut für Dynamik der Flugsysteme
D-8031 Weßling a.S.

Summary
An operational system is described which tracks moving objects in an image, and servoes a camera platform
to continue tracking. The system uses a fast hardware contrast evaluation logic supplying contour coordi-
nates to a supervisory control computer for further processing. Contrast evaluation is restricted to within
a rectangular window, which is adapted to the object size automatically. The mini-computer being free from
contour finding has enough time during one frame (20 msec) to process the contour data from the preceding
frame. Essential criteria in treating problem situations are the changes in contour from the present frame
to past frames after isolating the object motion from the superimposed camera motion. They are processed
in a kind of heuristic truth tables. Improvements of the tracker are achieved merely by software refine-
ments, that is by application of more elaborate algorithms.
Originally the tracker was designed for ground-to-air-tracking, meanwhile however ground-to-ground-tracking
was demonstrated for realistic situations too. Corresponding film material is shown at the conference.

1. Introduction
Imaging sensors are expected to become an increasingly important type of sensor for closed-loop controllers
in different kinds of "intelligent" systems such as occur in robotics or missile guidance. Automatic
tracking of moving objects with a computer-controlled TV-camera is one of the outstanding problems in this
area. Digitizing the video image and processing it then of course gives the greatest flexibility that can
be thought of; however systems of this kind so far are expensive and tend to consume a lot of computer
time; therefore real-time capabilities are very difficult to attain.

The operational system we are presenting here comes along without digitizing the pictures. Instead it uses
a contrast evaluation logic as a fast interface to the control computer so that real-time tracking is
ascertained with a high degree of reliability. The intelligence of the system is based on elaborate con-
trol programs that check the contour coordinates for plausibility and try to make reasonable decisions
in problem cases.

As is well known, the most challenging problem in tracking are occlusions. The target may occasionally be
either partially or totally occluded. The system presented here is to show that by elaborate computer
processing the occlusion problem can be handled in the non-digitized case, too.

2. The DFVLR TV-Tracker

2.1. The overall system
Figure 1 shows a simplified block diagram of the tracking system. The video signal is supplied by a black
and white gimballed camera and displayed on a color monitor. (The color functions of the monitor are used
to overlay the window location and threshold boundaries.) At the same time the analog video signal is com-
pared in the signal processing logic with an adjustable grey value level. Whenever the amplitude of the
video signal crosses this threshold, either rising or falling, the corresponding line number and pixel
number in this line are registered from counters synchronized with the frame and line sync pulses. At the
end of the line these values, having been accumulated in a fast memory, are transmitted to a process
computer.

Signal evaluation (from the hardware) occurs only within a window somewhat greater than the target exten-
sion. This window is normally adapted to the object size automatically by the computer, taking into
account the predicted target position to make sure the target never touches the window edge. The computer
then has the task of calculating the center and extensions of the target by the given contour values and
servoing the camera so as to null the deviation between object center and screen center.

The level 1 tracker is characterized by an enormous reduction of information flow. It uses only the first
and last threshold values in one line (inside the window) and it actually does not process all values of
the object contour, but only the extrema in the horizontal (x) and vertical (y) directions. So only four
values are actually encoded in each frame and used for decision making. These four extrema may be due to
the target contour, or they may stem from occlusions entering the window. It is the task of the logical
decision part of the computer program to cancel implausible values and replace them by estimated ones.

2.2. System Hardware
The main task of the signal processing logic is to register the contour and position of an object in the
TV picture and transfer this data to the computer. Except for the time required for data transfer, the
computer ist not involved in the image processing function.

The functioning of the signal processing unit is shown in Figure 2. Two versions of the system are illus-
trated in this figure. The level 1 system can handle only one object (threshold crossing) per line, whereas
the level 2 system can handle up to 32 objects. All internal signals operate synchronously with the TV
camera timing.

2.2.1. The Level 1 System
The position of any point inside the TV pitcture is described by its line number and pixel point number.
Consequently, the analog signal must be sampled at discrete times to provide a unique pixel point number
for each point on a scan line. For this, a crystal oscillator is used which provides pixel spacing equal
to line spacing. Taking into account the 3x4 aspect ratio of the camera, this results in a clock fre-
quency of 13 MHz. The pixel point counter is clocked by the output of this oscillator and reset by the
horizontal sync pulse.

In an analogous manner, the line counter is incremented by the horizontal sync pulse, and reset by the vertical sync pulse. Since both counters operate synchronously with the camera scan, the position of the electron beam is always known.

An object which differs significantly from the background in intensity results in a corresponding change in the analog voltage from the camera. This voltage is compared with a reference threshold, and whenever the two are equal, a comparator changes state. The threshold voltage can be fixed in advance by an operator or set to follow the average intensity of the window.

With the rising edge of the comparator signal, the contents of the pixel point counter are latched. This number then corresponds to the left boundary of the object. The right object boundary and corresponding line number are latched with the falling edge of the comparator output, and may be overwritten by successive falling edges in the same line. So only the last one is retained. The data are then transferred into buffer registers with the horizontal sync pulse, and a data transfer to the computer is initiated. By using these buffer registers, the computer is able to perform the read operation at any time during the next line.

When the new window position is computed, the positions of the four sides are transferred to a buffer register, and with the next vertical sync pulse, this data is transferred to the inputs of a digital comparator. This comparator compares the beam location (contents of the two counters) with the window coordinates, and permits an output from the analog comparator only if the beam is inside the window. With this signal from the digital comparator ("inhibit"), the window is marked on the monitor.

2.2.2. Level 2 System

As mentioned above, with the second version of the system, it is possible to record more than two threshold crossings, and consequently more than one object per line. The existing electronics were augmented by the addition of a first-in-first-out (FIFO) memory. The pixel point and line counters provide input to this memory. With the trigger signal from the analog comparator, the appropriate counter values are read into the FIFO and at the end of the line (horizontal sync pulse), data transfer to the computer is initiated. This data transfer is performed by a DMA channel which works by cycle stealing. Therefore, additional CPU time is not required even though significantly more data is available (as compared to level 1).

2.3. Structure of the camera control loop

In this section, the closed loop for camera control (in x and y) is briefly outlined in its dynamical behavior. An essential implication is the temporal relation between scene perception, processing and realization of camera motion. (table 1).

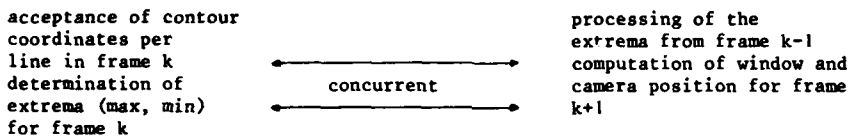| acceptance of contour coordinates per line in frame k determination of extrema (max, min) for frame k	concurrent	processing of the extrema from frame k-1 computation of window and camera position for frame k+1

Table 1. Software activities while frame k is recorded

With each frame starting pulse, the same operations are initiated, which is analogous to a normal sampled data system. While frame k is recorded from the video signal, the computer is processing the contour extrema of frame k-1. This program is concurrently interrupted by the transmission of the contour values of the running line of frame k. Inherent with this transmission is a min-max search so that at the end of frame k the corresponding extrema are known. Processing of frame k-1, being a background job is also finished at the end of frame k; the computer has by the end of frame k checked the suppositious target motion for plausibility, corrected it if necessary, predicted the target position two frames in advance, and calculated window coordinants and camera position for frame k+1.

Consequently, by the time data transmission from the window in frame k+1 starts, the camera has been commanded to move (by a command initiated after analyzing frame k-1), and will be in the correct position.

In a control theoretical sense, the plant here consists of a delay made up by two sampling periods (i.e. frame periods). In the z-transform domain often used with sampled data systems, a delay of 2 periods is characterized by a factor z^{-2}, while a prediction of two periods means multiplication by z^2 (see Figure 3).

As on the screen, only the relative position between target position and camera position is measured. To deal with occlusions, it is necessary to get the corrected target position by adding the camera motion. The problem solving and processing program (see Figure 3) tries to solve this task in "problem" cases too, and then predicts the ("true") target position two frames ahead in order to catch up the inherent delays. After processing occlusions and estimating velocities (if necessary), the effects of camera motion are reinserted to provide an error signal for the camera servo. A digital low pass filter processes this error signal first. This filter is charged with:

a) smoothing the small contour jumps from frame to frame which are due to the interlaced scanning techniques used in television,

b) supplying sufficient phase reserve ahead of the position-error integrator in the camera control circuit.

The dynamics of the closed loop are adjustable with the gain V (from Figure 3), so that the closed loop poles yield sufficient stability. As the system contains only one integration, an object moving with constant velocity generates a stationary position error that has no major influence on the tracking capability.

2.4. Treatment of problem cases (e.g. occlusions)

The most critical part of the control program has the task of checking the measured (apparent) object extrema for plausibility. For this purpose it uses:

a) the past, digitally filtered speed values of the contour extrema xmax/min, ymax/min and the computed overall speed of these in x and y.

b) the object extension averaged over several frames.

Processing a frame implies checks for two kinds of disturbances. (See Figure 4).

a) a jump disturbance, which occurs when suddenly one or several contour extrema (in the following abbreviated as CE) coincide with window edges; in this case it can be assumed that an occlusion has entered the window, and that part of the occlusion overlaps the edge of the window of course, for this to work, the real target size never should exceed the edges of the window. The window size is continuosly monitored to insure that this assumption is true.

b) a velocity disturbance, which occurs when the measured interframe change of a contour extremum (CE) does not have the same sign as the computed overall velocity.

Taking account of the typical structures of occlusions (trees, poles, horizon), different algorithms were developed for the x and y directions, which in condensed form, are as follows:

x-direction
In case of a jump disturbance or a velocity disturbance the velocity of the disturbed CE is not updated; if both CEs are disturbed, pure prediction is performed, i.e., window size and speed keep their old values. Moreover, in the case of a velocity disturbance in one CE, the disturbed CE is estimated from the latest object size and the other (undisturbed) CE. Object size is updated only in undisturbed cases.

y-direction
Here only jump disturbances are checked. When a disturbance is recognized, the corresponding CE is again estimated from the undisturbed one and updating of CE velocity is omitted. If both CEs are disturbed, prediction is used.

If the growth of the y-extension taken over several frames exceeds a threshold, then that CE which is the leading one with respect to object velocity is used to briefly (i.e. for one frame) draw after it the other CE. Then (in the next frame) it will be determined whether the object size actually increased so much or whether it must overcome an occlusion stationary with respect to y.

For example, in Figure 4, jump disturbances and velocity disturbances are sketched for the x-coordinant direction. The jump disturbance at the left CE indicates that an occlusion has entered the window- presumming of course, that the target does not touch the window edges. So for velocity updating, this CE is no longer used but tries to move to the end of the occlusion and to "wait" for the target there. Similar jump disturbances in y however would immediately fix the window in y. This scheme takes account of the fact that in many practical cases nearly horizontal movements coincide with vertical structured occlusions, for example, trees and buildings.

Registration of velocity disturbances as indicated in Figure 4 assures that indepent of whether the target reverses its direction behind an occlusion, the window is finally detached from the occlusion.

With level 2 hardware as described in chapter 2.2. all threshold crossings within the window are registrated. For the tracking purposes the software remain nearly as described, however if there are more than one threshold crossings in a line the run-lenght data are checked for plausibility and those segments are rejected, the center of which is not within the predicted contour extrema. This improves performance especially in case of oblique occlusions entering the window.

Though the tracker initially was designed only for ground-air tracking with good contrast, it proved in field experiments to show good tracking capabilities in ground-ground tracking with occlusion situations too. Thus it can be shown that by intelligent processing of dramatically reduced information (4 values per frame) needing only 30 - 40 % of real time on the computer, surprisingly good results could be achieved.

3. Conclusion
The system presented here shows that by implying elaborate computer programs it is possible to achieve good tracking capabilities although the TV-image is not digitized. Nevertheless it must be stated that the algorithms are "synthetic", i.e. are based on ideas of how we intuitively perceive that a computer could be directed to perform the tasks in question. No learning capability is provided. Thus in future our main interest will be in how to make machines - like the one presented here - change their behaviour from experience made in the past, i.e. introduce a learning phase in which a set of typical situations is shown to the tracker. Again this problem is to a great extent independent from whether a picture is digitized or not.

References
Martin, W.,and Aggarwal, J., "SURVEY, Dynamic Szene Analysis"
    Computer Graphics and Image Processing, Vol 7 no 3, June, 1978.

Martin, W.,and Aggarwal, J., "Dynamic Scene Analysis: the study of Moving Images"
    TR No. 184, Information System Research Lab, Electronics Research Center,
    University of Texas at Austin, 1977.

Nagel, H., "Analysis Techniques for Image Sequences", 4IJCPR, Kyoto, Japan, 1978.

Roach, J., and Aggarwal, J., "Computer Tracking of Objects Moving in Space"
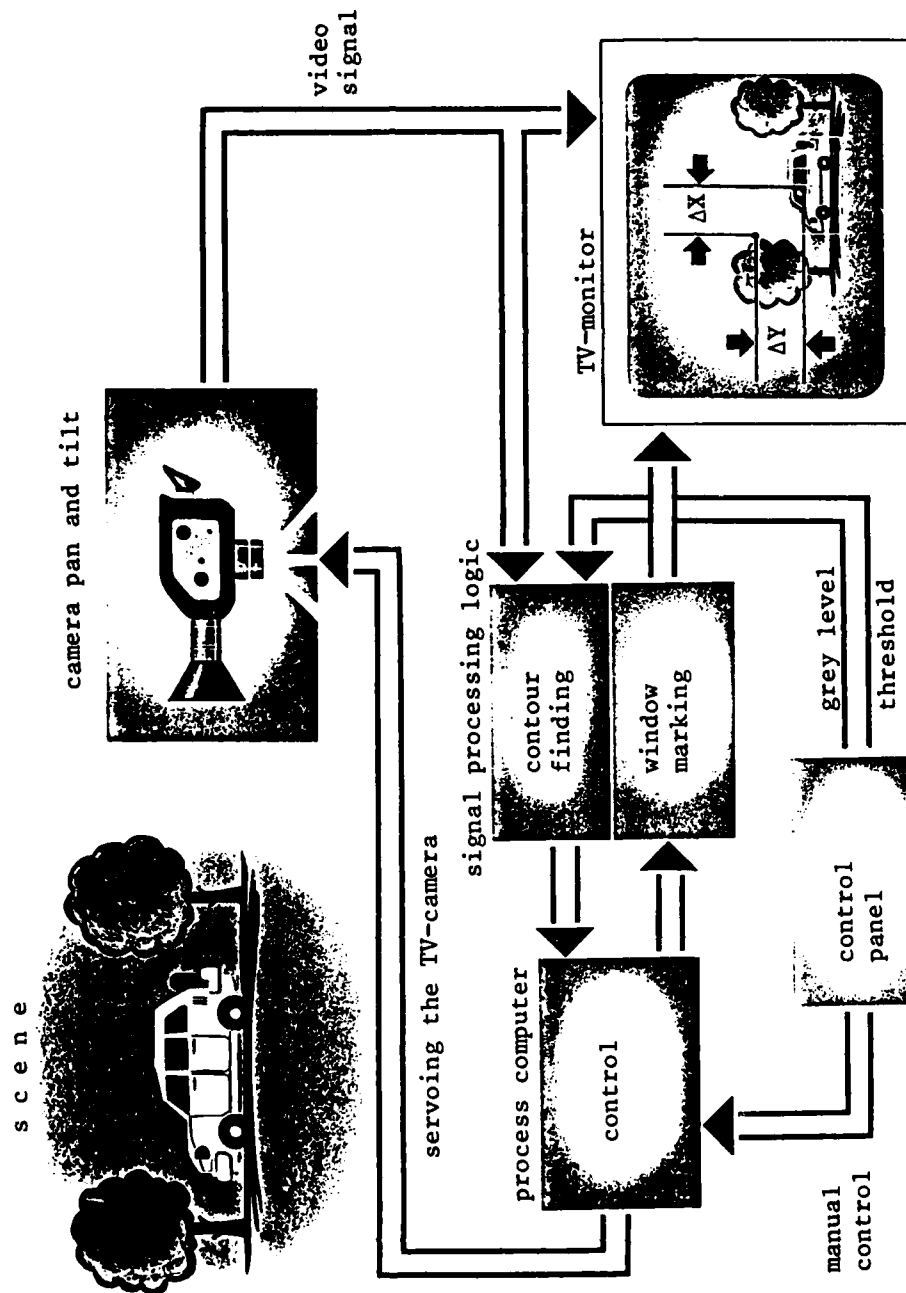   IEEE Trans on PAMI, 1-1, No. 2, 1979.
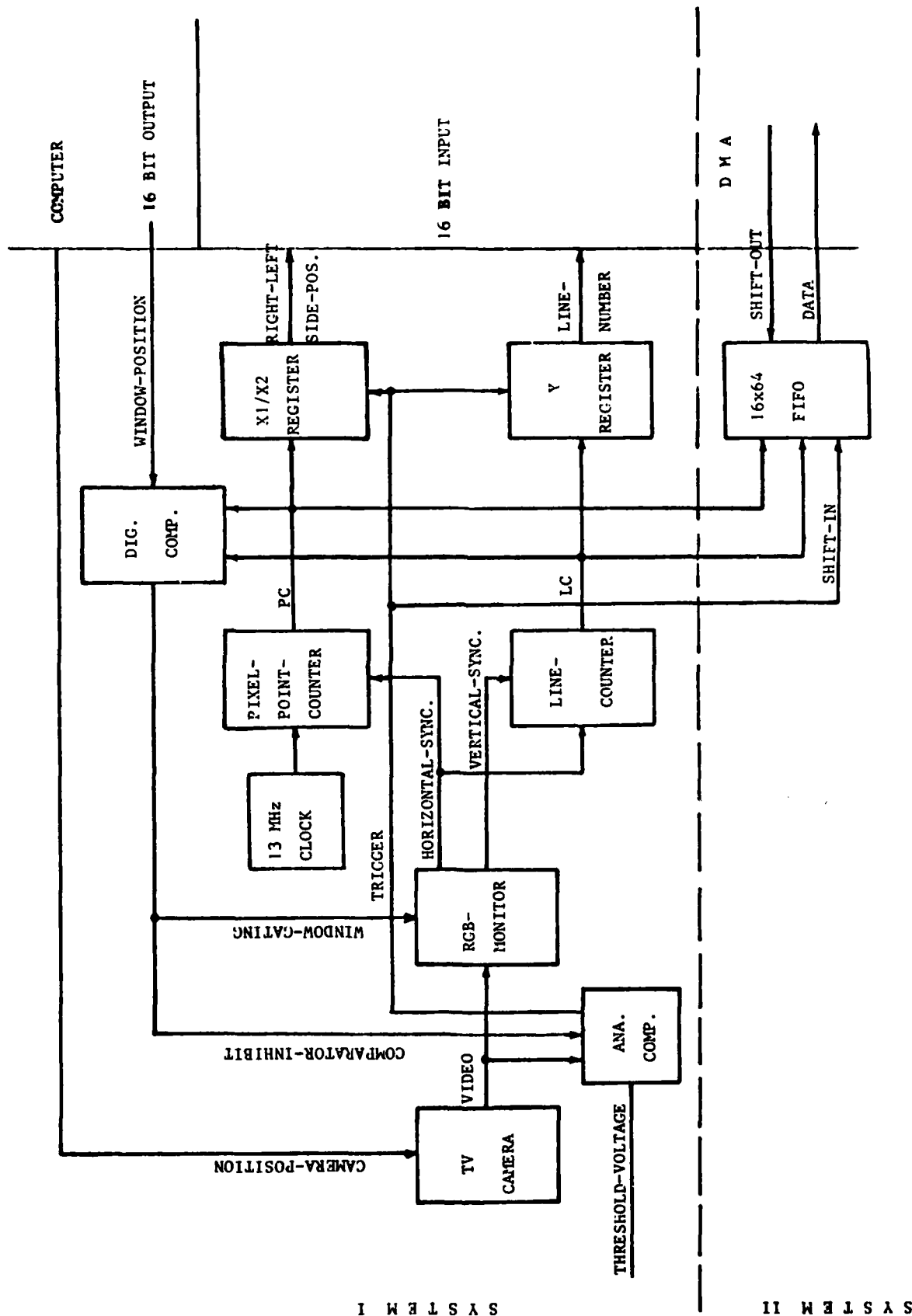
Fig.1 Block structure of the tracking system

Fig.2 System hardware

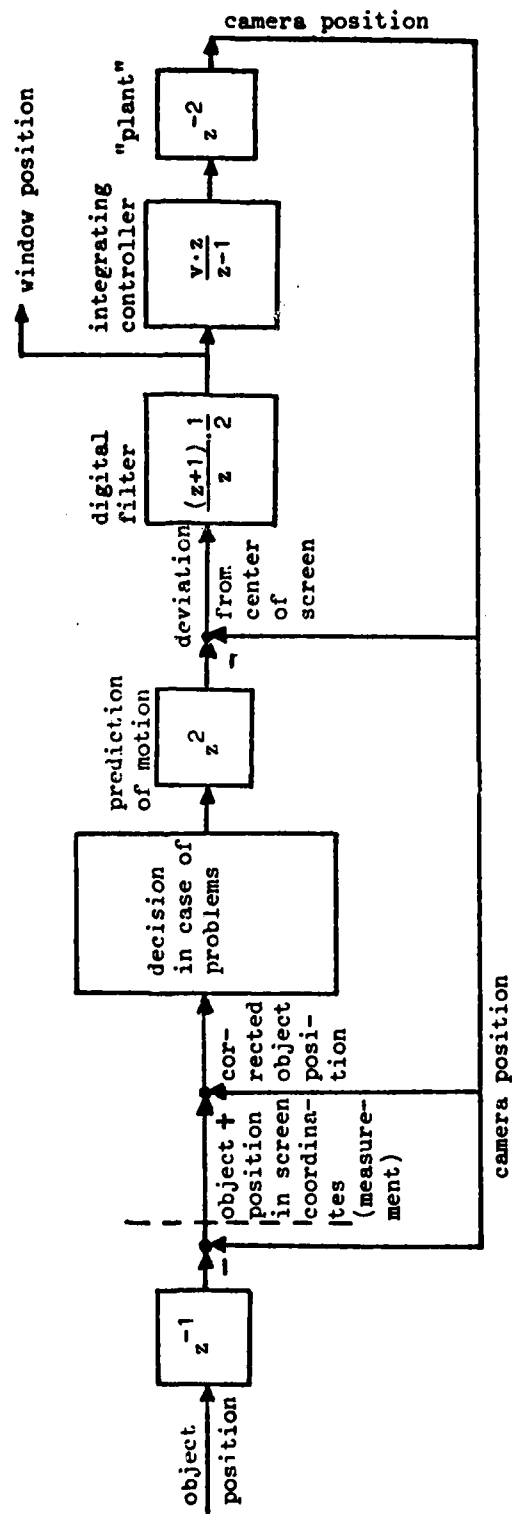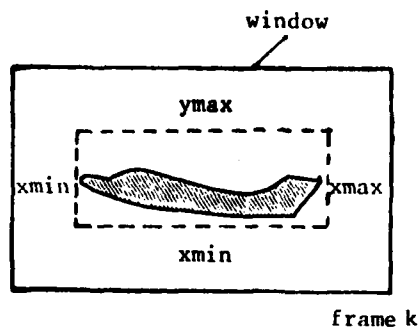Fig.3   Structure of the camera control loop

Figure 4a

normal case with contour
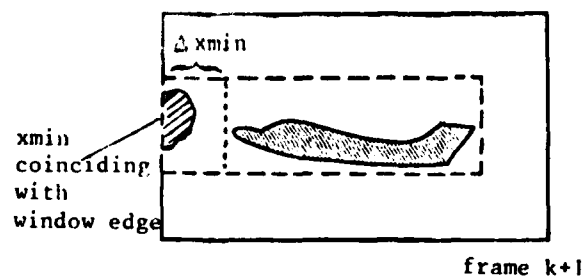extrema (CE) xmin, xmax
ymin, ymax

Figure 4b

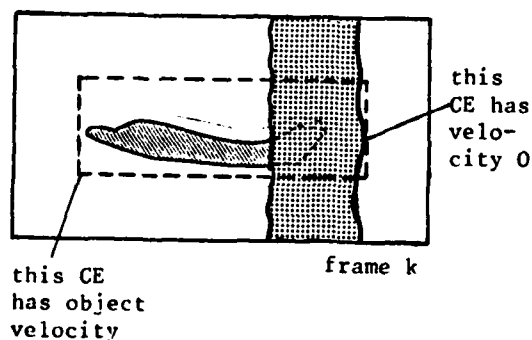jump disturbance with
any CE coinciding with
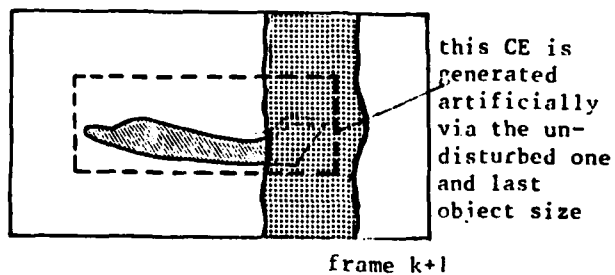a window edge

Figure 4c

velocity disturbance

Figure 4d

detaching   from occlusion
by estimating disturbed CE
from the undisturbed one

Fig.4   Problem cases

POURSUITE NUMERIQUE DE CIBLES

Michel GIRARD

Etablissement Technique Central de l'Armement

94114 ARCUEIL CEDEX   FRANCE

RESUME

La poursuite de cibles que nous avons étudiée s'appuie sur une méthode de corrélation logique d'images.

La cible formée de quelques dizaines de pixels peut être caractérisée par des points situés aux angles de frontières importantes de densité.

La majorité de ces points sont présents dans des images successives et ont le même vecteur de translation. Ils sont donc différenciés du fond.

Par cette méthode, la corrélation est rapide, efficace, même sur fond perturbé. Elle est aussi "intelligente" car l'évolution du nombre de points permet souvent d'en connaître la cause et de choisir l'algorithme le mieux adapté pour essayer d'éviter un décrochage éventuel de la poursuite.

INTRODUCTION

Le but de l'étude a été de réaliser un système permettant de corréler deux images de façon automatique, le couple d'images pouvant être quelconque. Mais très rapidement, on s'est aperçu que la méthode utilisée était très intéressante dans le cadre de la poursuite de cible de part sa rapidité, sa qualité dans le résultat et son "intelligence" dans la séparation de la cible et du fond.

Le principe fondamental repose sur une corrélation logique de deux images successives. Les deux images sont réduites à deux listes de points caractéristiques de coordonnées (x,y) prenant des valeurs fonction du contexte local. La comparaison de ces deux listes nous permet de créer deux classes, voire plus, la principale correspond à la cible, la seconde au fond (ou a une deuxième cible en cas de poursuite multicibles), les autres classes n'étant pas significatives.

Les points caractéristiques peuvent évoluer dans le temps. Un test de confiance doit être appliqué afin de savoir si :

- la corrélation est considérée comme bonne

- le résultat n'évolue pas vers une solution de décrochage auquel cas, il est nécessaire de passer en mode mémoire.

Dans ce document sont décrits :

- Les principes généraux de la méthode utilisée

- L'application de ces principes dans le cadre de la poursuite de cibles

- Le matériel "poursuite"

I - <u>PRINCIPES GENERAUX DE LA METHODE</u>

    I-1. <u>Généralités</u>

        Nous nous proposons d'associer par couples les points représentatifs de l'objet. Ces points seront dits "points caractéristiques".

        L'idée directrice de la méthode est basée sur l'hypothèse que l'oeil détecte dans un premier stade les lignes séparant des surfaces de teintes différentes et en particulier les sommets des angles formés par ces lignes.

Hypothèse : les deux photos ou deux parties homologues de la photo se correspondent par une translation ou de faibles déformations (inférieures au pas d'analyse).

        Dans une numérisation, la taille du spot étant supérieure à la distance entre deux positions consécutives, une frontière est définie par trois points :
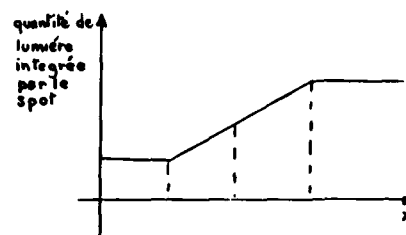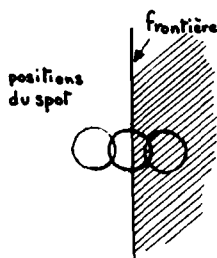
        Un point caractéristique peut être défini par neuf points. Le point lui-même et les huit qui l'entourent.

        Nous allons essayer d'extraire les points caractéristiques de la photo en faisant des considérations locales sur ces 9 points.

La méthode comporte deux parties :

    - extraction des points caractéristiques,

    - corrélation.

    I-2. <u>Recherche des points caractéristiques</u>

        Nous avons défini les points caractéristiques comme des points de rupture sur une frontière.

        Ces points doivent donc vérifier deux conditions :

        - se trouver sur une frontière, donc la variance du point doit être forte,

        - la frontière doit former un angle en ce point.

        Un point caractéristique peut se présenter de plusieurs façons :
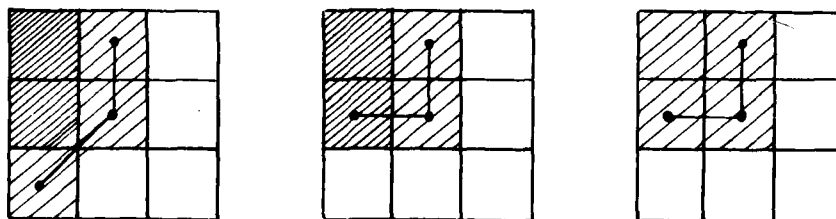
        Si un point est caractéristique, il se peut qu'un ou plusieurs de ses voisins le soit également.

Dans la figure ci-contre, 4 points peuvent être caractéristiques suivant la définition précédente.
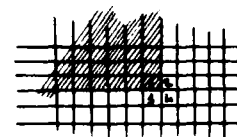
Problème : Soit un point de l'image et ses 8 voisins.

Ce point est-il caractéristique ?

Il faut trouver un opérateur qui puisse répondre à cette question.

On peut vérifier les performances de l'opérateur utilisé d'une façon très simple : on détermine sur une photographie ou sur une restitution sur imprimante les points qui à l'oeil semblent les plus caractéristiques. Si l'opérateur retrouve une grande partie de ces points, il peut être considéré comme valable.

Choix de l'opérateur

1. Si le point est caractéristique, il doit se trouver sur une frontière. Nous allons calculer la variance du point :

$$v' = \sum_{i=1,3} \sum_{j=1,3} (T'_{ij} - m')^2$$

avec $m' = \frac{1}{9} \sum_i \sum_j T'_{ij}$

Les T'_{ij} étant les nouvelles transparences après normalisation.

2. La frontière doit former un angle en ce point.

C'est la partie la plus délicate de cette partie du programme. Il est difficile de trouver un algorithme simple qui tienne compte de toutes les possibilités. On est aussi parfois incapable de prendre une décision.

Le point central est-il caractéristique ?

De tels cas peuvent se produire si l'ordre de grandeur des détails de la photo est le pas de numérisation et aussi s'il existe des points aberrants.

D'après la partie 1 la frontière existe et passe par le point central.

Il faut déterminer pour les 8 points qui l'entourent les 2 qui ont la plus forte probabilité de se trouver sur la frontière.

Nous allons utiliser un codage de Freeman pour représenter le carré 3 x 3 :

3	2	1
4	C	0
5	6	7

Dire que le point d'indice i se trouve sur une frontière cela signifie que la transparence en ce point est assez proche de celle du point central et que la pente est forte.

Nous allons donner deux définitions : i $\in$ {0,7}

Soit $T(i)$ la transparence au point i et T_c la transparence au centre.

On appelle gradient circulaire le nombre $G(i) = T(i + 1) - T(i - 1)$; $i + 1$ et $i - 1$ étant calculés modulo 8.

On appelle gradient par rapport au centre le nombre $g(i) = T(i) - T_c$

D'où :

$$G(1) = T(2) - T(0)$$
$$G(7) = T(0) - T(6)$$
$$G(0) = T(1) - T(7)$$
$$\text{et} \quad g(1) = T(1) - T_c$$

<u>Définition</u> : un point du pourtour sera dit point frontière si :

- $|G(i)| > |g(i)|$

- il est un des *extrema de la fonction K(i) telle que :*

- $K(i) = \left[\big| |G(i)| - |g(i)| \big| \right] \times \text{signe} \left[G(i) \right]$
  si $|G(i)| > |g(i)|$

- $K(i) = 0$ si $|G(i)| < |g(i)|$

Nous allons expliciter cette définition et donner quelques exemples.

Considérons un angle droit numérisé. Du fait du diamètre non négligeable du spot , la réponse numérique obtenue pourra être fort différente.
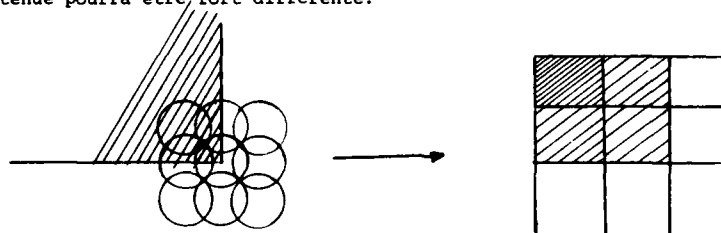
Les points d'indice 2 et 4 auront des valeurs comprises entre T_{min} et T_{max} mais ces valeurs dépendront de la position du spot par rapport aux côtés des angles. Par contre les valeurs de T_{min} et T_{max} seront *constantes à très peu près.*

Si l'on veut estimer une pente au point d'indice 2, il est préférable de prendre la valeur T(3)-T(1) plutôt que $|T(2) - T(1)| + |T(3) - T(2)|$, par exemple. Cette particularité explique le choix de la dé-finition du gradient circulaire.

Il faut aussi tenir compte de la transparence du point central (T_c).

Considérons les exemples suivants :

Dans le premier cas, il semble logique de faire passer la frontière par les points 2 et 4 et dans le *deuxième cas par les points 1 et 5.*

On va utiliser le gradient avec le centre, g , pour tenir compte de ce phénomène.

La première partie de la définition $|G(i)| > |g(i)|$ impose au point frontière d'avoir une valeur plus proche de celle du centre que de celles de ses voisins. Dans le cas des exemples, le point 3 en par-ticulier ne doit pas être choisi.

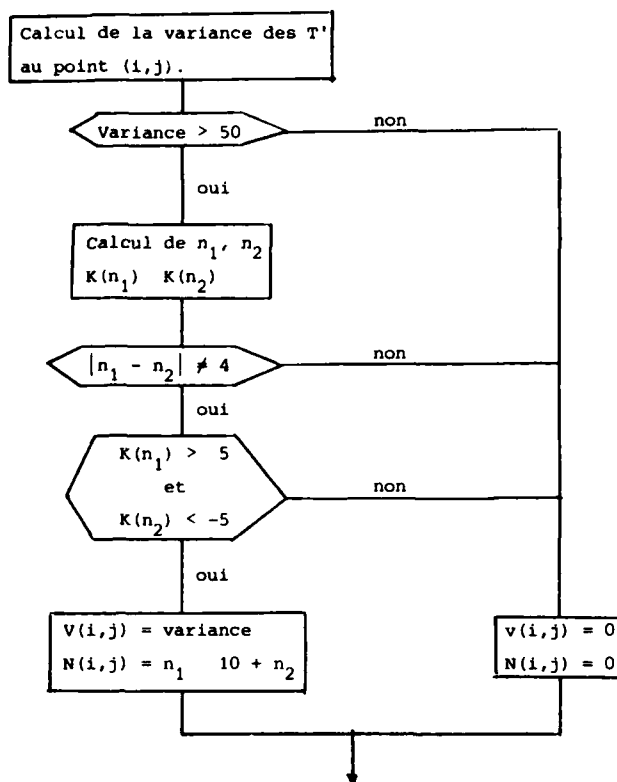$|G(i)| - |g(i)|$ est toujours positive ou nulle. Le fait du multiplier par le signe de G(i) revient à conserver le sens de la pente et donc de trouver les deux points frontières.

Un point sera caractérisé par 2 nombres n_1 et n_2 qui sont les directions des vecteurs de Freeman de la frontière. Pour que le point soit caractéristique, il suffit que ces deux nombres soient diffé-rents de 4.

$$|n_1 - n_2| \neq 4$$

Récapitulation

Calcul de la variance des T'
au point (i,j).

Variance > 50 — non

oui

Calcul de n_1, n_2
$K(n_1)$ $K(n_2)$

$|n_1 - n_2| \neq 4$ — non

oui

$K(n_1) > 5$
et
$K(n_2) < -5$ — non

oui

V(i,j) = variance
N(i,j) = n_1 10 + n_2

v(i,j) = 0
N(i,j) = 0

A ce niveau du programme, nous possédons deux tableaux. Le premier V(i,j) contient les variances et le deuxième les nombres caractéristiques des frontières. On utilise la forme N(i,j)= n_1 10 + n_2, ce qui permet un gain de place en mémoire.

Nous avons vu précédemment que si un point était caractéristique, certains de ses 8 voisins le seront aussi. Pour la suite du programme, il est nécessaire d'éliminer ce genre de situation. Pour cela, nous allons prendre les points du tableau qui sont des maximas locaux pour la fonction variance.

II - LE PROGRAMME POURSUITE

II-1. Généralités

Nous utilisons actuellement, pour réaliser les acquisitions de données, la caméra RETICON RA 50X50 contenant 2500 photodiodes accompagnée de son électronique de commande.

Les données obtenues à l'aide de cette caméra sont enregistrées sur bande magnétique image par image. Cette phase, qui ne nous permet pas de travailler en temps réel, n'a pas été supprimée pour les raisons suivantes :

- la caméra n'est pas montée sur support amovible, donc le traitement total en temps réel n'est pas possible,

- il est plus aisé d'utiliser des bandes magnétiques pour tester les programmes que de faire des acquisitions à chaque passage,

- l'interface caméra-unité centrale d'ordinateur ne pose aucun problème et son étude n'a pas été le but principal de ces travaux.

Les données sont traitées par le programme POURSUITE. Ce dernier recherche les éléments de l'image permettant de "reconnaitre" l'objet en mouvement.

II-2. Déroulement du programme POURSUITE

a) Programme Principal - P.P

Il permet d'appeler les sous-programmes fonctionnels. D'après les résultats obtenus par CORRE, il calcule le déplacement de l'objet et les coordonnées du point central de l'objet dans l'image suivante, coordonnées fournies par la suite à PCAR.

b) GRAD

Ce sous-programme calcule le gradient et les vecteurs de Freeman en chaque point de l'image.

c) PCAR

Ce sous-programme établit une liste de n premiers points caractéristiques trouvés par recherche circulaire autour du point central de l'objet à suivre.

Cette méthode de recherche des points a été choisie arbitrairement. Il est possible, en fonction des besoins utilisateurs, de créer des sous-programmes optionnels PCAR : PCAR1, PCAR2, ..., basés sur d'autres méthodes de recherche des points caractéristiques.

d) CORRE

Parmi tous les couples de points caractéristiques obtenus pour deux images successives, CORRE décide quel est l'ensemble de ces couples appartenant à l'objet, décision prise en considérant les translations entre les points de chaque couple :

A l'entrée de ce module les deux photos à traiter sont représentées par deux listes de points caractéristiques. Ils sont définis par les coordonnées et la valeur des vecteurs orientation.

L'élément fondamental d'un point caractéristique est la valeur du vecteur orientation.

La première étape de calcul consiste en la recherche pour chaque point caractéristique de la première photo, des points caractéristiques de la deuxième photo qui ont l'orientation la plus proche.

Pour chaque couple ainsi défini, on détermine la valeur du vecteur de translation qui amène le point de la première photo sur son homologue de la deuxième ; on explore ainsi tous les points caractéristiques de la première photo et on dresse le tableau des vecteurs de translation avec pour chacun le nombre de couples de points qui se correspondent suivant celui-ci.

Chaque classe regroupe les vecteurs s'écartant de $\pm$ 2 pas d'analyse (suivant les coordonnées x et y) autour du vecteur moyen, celui-ci étant recalculé chaque fois qu'un vecteur nouveau entre dans la classe.

Ce qui permet, si les premiers vecteurs de translation s'écartent trop du vecteur translation moyenne, de ne pas créer de classes artificielles.

Chaque point pouvant avoir plusieurs homologues, la somme des éléments des classes de vecteurs de translation est en général supérieure au nombre de points caractéristiques.

On élimine arbitrairement les classes ayant moins de quatre éléments en les supposant non représentatives du mouvement d'un objet. Si sur le raster de 50 x 50 il y a plusieurs objets en mouvement indépendants, ce seuil serait à reconsidérer.

On range alors les classes de vecteurs translation en fonction décroissante du nombre de leurs représentants. Sur les films que nous avons traités, on obtenait en général deux classes, exceptionnellement trois.

Dans les cas les plus fréquemment rencontrés, la première classe peut être considérée représentative de l'objet, la seconde classe représentant le paysage.

L'évolution du nombre de points des classes d'une image à l'autre peut nous permettre de connaitre, dans certains cas, les causes d'un décrochage éventuel de la poursuite et de remédier au mieux à cette défaillance.

Dans sa version de base, le sous-programme de corrélation fournit au programme principal la translation (ici en nombre de photodiodes) que l'objet a subi entre les deux images. Cette donnée permet de calculer :

- le déplacement de l'objet dans l'image,

- la position de l'objet dans l'image suivante. Cela se traduit par une inertie de la caméra. Les coordonnées de son centre seront fournies à PCAR dans la passe suivante pour la recherche circulaire des points caractéristiques.

- le déplacement, la vitesse et l'accélération tangentiels de l'objet dans la scène.

Déplacement réel de l'objet (sa vitesse, son accélération)

Si n est le déplacement observé au niveau de l'image, alors celui qu'a subi l'objet sur la scène est :

$$TR = \frac{npD}{f} \quad \text{avec}$$

p : distance centre à centre de deux photodiodes voisines

D : distance objet-caméra

f : focale de l'objectif

Dans le cas présent :

$$f = 12,5 \text{ mm et } p = 0,0975$$

Ainsi :

$$\boxed{TR = 0,0078 \; n.D}$$

L'image étant sous une forme discrète, nous avons une erreur sur la mesure de la translation.

L'erreur relative est :

$$\frac{\Delta TR}{TR} = \frac{1}{n}$$

Entre deux photos non successives (en ajoutant les translations calculées au cours de chaque passe), l'erreur relative est au maximum égale à :

$$\frac{\Delta TR}{TR} = \frac{\ell}{n}$$

où ℓ est la largeur (en nombre de pas) de l'objet dans l'image.

En effet, en utilisant cet algorithme de corrélation, nous ne pouvons reconnaître la forme de l'image. Les points corrélés à l'instant t ne sont donc pas nécessairement les mêmes que ceux corrélés à l'instant t' (sauf si t' - t = t_0 est le temps entre deux acquisitions successives). Il peut exister un "glissement" des points corrélés dans l'objet, glissement limité à sa largeur.

D'après les résultats obtenus jusqu'à maintenant, nous n'avons pas observé de glissement important et :

$$\frac{\Delta TR}{TR} \approx \frac{1}{n}$$

La vitesse est déterminée si l'on connaît le temps entre deux acquisitions. L'erreur relative sur la vitesse sera de :

$$\frac{\Delta V}{V} \approx \frac{1}{n}$$

De même pour l'accélération :

$$\frac{\Delta \gamma}{\gamma} \approx \frac{\Delta V}{V} \times \frac{V_2 + V_1}{V_2 - V_1}$$

III - LE MATERIEL "POURSUITE"

III-1. Généralités

Les résultats obtenus à l'aide de l'algorithme "POURSUITE" ont été satisfaisants. Cependant, le temps d'exécution sur l'ordinateur de l'E.T.C.A. (CIMSA, Mitra 15) est loin du temps réel : 1,5 seconde par image de 50 x 50 points en utilisant deux listes de 20 points caractéristiques chacune.

Nous avons entrepris de réaliser une maquette basée sur l'algorithme précédemment décrit.

La maquette peut être scindée en deux corps indépendants :

- lecture de l'image, détection et listage des points caractéristiques,

- corrélation des points caractéristiques et fourniture des résultats.

Seule la première partie est en cours de réalisation.

III-2. Principe

L'image originale provenant de la caméra RETICON est codée sur 8 bits. Des corrections d'offsett permettent de s'affranchir des décalages des tensions de référence entre chaque photodiode.

La recherche des points caractéristiques se fait sur une image binaire obtenue par seuillage de l'image des variances afin de simplifier le câblage. Pour réaliser une telle opération sans risque d'erreurs, il est nécessaire,avant de seuiller l'image, de la lisser afin d'éliminer tout bruit.

L'algorithme, pour obtenir les points caractéristiques, se déroule donc schématiquement de la façon suivante :

- Calcul de la variance,

- Lissage par zone par rapport à la variance moyenne locale,

- Seuillage,

- Détection des points caractéristiques.

Le choix de la taille des zones se fait en fonction du nombre de points caractéristiques donnés, du nombre de transitions de l'image lissée, du nombre de points caractéristiques obtenus dans l'image précédente.

Le seuil est fonction des variances moyennes locales de chaque zone.

Pour détecter les points caractéristiques, chaque point peut être codé sur 8 bits en fonction de son voisinage :

ex :

	1	2	3	
	1	0	0	
8	1	1	0	4
	1	1	0	
	7	6	5	

n° du bit	1	2	3	4	5	6	7	8
	1	0	0	0	0	1	1	1

Il suffit de comparer chaque mot aux mots d'une table listant tous les points caractéristiques possibles.

Le résultat va être codé sur un mot de 8 bits composé de 0 sauf dans les deux directions privilégiées où nous y trouvons des 1.

IV - CONCLUSION

Cette étude a permis de mettre au point un algorithme simple de poursuite de cibles sur fond perturbé, et de cabler les parties de programme les plus pénalisantes en temps.

La validation de la méthode a été faite en utilisant :

- des films que nous avons numérisés vue par vue

- un objet maquette en salle.

Ces essais ont montré que le système fonctionnait de façon satisfaisante, qu'il est très prometteur quant à son avenir et son utilisation.

TARGET TRACKING USING AREA CORRELATION

R.M.B. Jackson
EMI Electronics Ltd
Victoria Road
Feltham, Middx

SUMMARY

This paper describes an Image Tracking System which uses an area correlation technique.   The system operation and method of implementation are described and a short film of a tracker operating is shown.

INTRODUCTION

With the increasing use of electro-optic imagers in weapon systems for aircraft, there is a need to provide an automatic track of targets of interest to relieve the operator – who may be the pilot – of this task.   These systems must provide an output of the line of sight to the target, relative to an air frame datum, with the necessary precision and band width to allow weapon aiming, command and control.

The method described here is designed to operate on video which provides a grey scale image of the spatial pattern of the target.   This pattern may come from either a TV (daylight or low light level) or an Infra Red (IR) imager.   The spectrum used is not important to the tracker provided a spatial and grey level pattern is available.

The process examines the match of a stored area reference pattern with the patterns present in successive fields of video, applying equal weights to the comparison of equal areas.   This provides considerable immunity to effects of small changes in the pattern which may be caused by smoke, fluttering flags, branches, small obscurations from ground, and so on.

For the purpose of this paper I will confine the description to a tracker which uses video provided in TV format, because this is well known and is commonly used for display purposes for IR as well as TV imagers.   It is of course possible to design the tracker to operate with any format providing successive fields or frames of video have a known relationship to one another.

VIDEO SAMPLING

These systems use digital processing of the video waveform so the first process is to sample the video and digitise the amplitude to provide a digital representation of the pattern.   This process is shown in Figure 1.   The video comes from the imager in two interlaced fields of horizontal lines.   The odd and even fields are shown in the 'magnified' circle.

The tracker operates in each field separately.   This avoids storing one field to be combined with the next and also increases the rate at which measurements of target positions can be provided as an output.

Each line is sampled along its length and the grey level or amplitude of each sample is digitised.

The number of samples taken along the line is matched to the horizontal resolution available from the imager.   We normally take two samples for each cycle of resolution available to preserve the image data.   A typical number – for normal TV bandwidth – is 512 samples per line.

The grey level is digitised to six bits providing 64 grey levels.   This is usually sufficient for the type of video obtained from FLIR or TV imagers.   It could however be increased if the signal to noise ratio of the imager warranted the resulting increase in hardware and power.   Some earlier machines used fewer levels – typically three or four bits.

TARGET DESIGNATION

The next step is to designate the target to the tracker.   We ask the operator to do this.   The format of the display is shown in Figure 2.

Using a joystick control he moves the small box and crosswire marker over the target of interest, adjusts the box for an approximate fit to the target and selects automatic track.

The aim point he has selected may be adjusted while the machine is tracking – without losing lock – if he wishes to refine its position later in the engagement.

The size of the box may also be readjusted during automatic tracking if he wishes.

AREA CORRELATION

The machine now stores from the video at the time auto track is selected the pattern contained in the area of the field of view designated by the small box.

This is stored in the Inner Patch or Target Data store and is used as the memory of the target pattern for the tracking process. As we have seen this is a three dimensional pattern of spatial and grey level descriptors of the target. This is shown in Figure 3.

The size of the target patch has been selected to suit the size of the target pattern. However it is not necessary to fit it precisely to the target outline and some background in the patch will not cause loss of lock. This is because of the area averaging effects of the correlation process.

It is not necessary either, to include the edges of the target within the patch. If there is a spatial and grey level pattern resolved within the target outline this can be tracked and this may well be the case if the target image is large.

Typical sizes used for the Target Data patch range from 9 x 9 to 23 x 23 samples. It may of course be rectangular rather than square if required.

On succeeding fields of video information another store - the Outer Patch store is loaded with data surrounding the last known position of the target. This is also shown in Figure 3.

The size of this store is selected in the design to suit the dynamics of the system requirements. It must be large enough to be sure that the target will not move out of this area of the field of view within one field time of the imager. This movement may be caused either by movement of the target within the background or of course by movement of the boresight of the imager due to aircraft movement, vibration and so on.

A search radius is usually defined from these considerations and the store is designed to allow the largest Inner Patch to excurse over the area defined by this radius. A typical figure for a low flying aircraft attacking tank targets is a search radius of 10 samples.

The correlation process is now carried out over the search area by moving the Inner Patch one sample at a time to each possible position within the Outer Patch. This is shown in Figure 4. At each position a score for the match of each sample in the Inner Patch with the corresponding sample in the Outer Patch is developed. These scores are accumulated over the Inner Patch area and this figure is allocated to the sample address corresponding to the centre of the Inner Patch store.

The scores so generated are collected and may be plotted as shown in Figure 5.

They form a surface which has a peak at the measured position of the target. This position can be expressed as an address in the sampling matrix.

TARGET POSITION INFORMATION

This target position is required by the aiming and control systems referenced to a datum. It is often convenient for this datum to be the centre of the field of view (see Figure 2), but it may be a point at another position related to the field of view.

The target position is then expressed as an x, y, co-ordinate in picture samples from this datum.

This position information is available in each field of video with a delay, for processing, from the instant when the Outer Patch store is completely loaded. This delay is fixed by the number of calculations which have to be carried out, which depends on patch size and search area. However it can be reduced by providing parallel chains of arithmetic logic, so there is a trade-off of delay and hardware complexity size and power. Many servo systems require - or at least prefer - a fixed value for this delay, so it is often set at the maximum for a particular tracker system, and answers from the smaller patch sizes are delayed and output at this time.

CHANGES IN TARGET PATTERN

As the target moves through the background, or as the aircraft approaches, the image of the target will change. We must provide an automatic means of coping with these changes.

The effect on the correlation system as the image changes will be to reduce the correlation score (see Figure 5). Because of the area nature of the process small changes will produce very small reductions in peak score. But larger changes such as change of orientation or magnification due to approaching the target will result in significant reductions in peak score.

These changes can be detected by monitoring the peak score and by setting a threshold. If the peak score drops below this threshold the machine triggers an 'update' where the contents of the Inner Patch store are discarded and a new target image is loaded by taking data centred on the latest target position. This is then used for subsequent fields as before.

The threshold is set so that the update is triggered before loss of lock occurs, but sufficiently low to ensure that it does not occur too often. This is to reduce the effects of the one sample uncertainty caused by the granularity of the target position measurement leading to a significant shift of aim point. The effects which lead to the requirement for an update usually occur relatively slowly compared with the field rate of the imaging system so this is not normally a problem.

EFFECTS OF OBSCURATION

If the target is obscured so that there is no target image the tracker cannot track. This may occur through natural or man·made objects in the scene coming into the line of sight to the target or through deliberate deployment of countermeasures such as smoke.

The method of overcoming these effects is to use a track memory of previous results and predict the movement of the target past the obstruction. Two and three term tracking filters are used for this purpose and these can provide considerable protection against these effects.

The predictive mode is triggered by examining the correlation surface peak score and shape of the peak. Often the reduction in score because of an obstruction is rapid compared with the signature change due to orientation and so on. It is also usually much more extensive. These two aspects are used to differentiate between the two cases.

When obscuration is detected the tracker continues to position the search area on successive fields centred on the predicted position of the target. The update procedure is inhibited and the correlation search is carried out as usual. When the target is reacquired tracking proceeds normally.

During the obscuration the target position output is giving the predicted position to allow the servo control to continue. However a signal is available to other systems to indicate that this is predicted and not tracking information.

This can provide information for a short period. If the obscuration continues for a long period however lock will be lost.

PERFORMANCE

The performance of these systems depends on the image quality provided by the imager, the signature presented by the target and the consistency of this information as it passes through the field of view.

Video signal to noise ratio is important, however the area averaging, which occurs in the area correlation process, reduces the effect of noise spikes on the correlation score. In very severe noise the surface itself becomes noisy and it is possible for a false peak to be detected.

Spatial frequency response and grey level response - related to imager M.T.F. - are also important. The sampling process is designed to use the maximum spatial frequency information available and this also sets the measurement grid for target position information. The peak of the surface is detected to the nearest sample and this sets the accuracy of the target position measurement. The sharpness of the correlation surface peak is a function both of the spatial frequency content and the grey level content in the target pattern. If the peak is flat a smaller noise effect could cause a shift of the peak detected.

It is important that the imager response to a particular signature is consistent, at least through the search area of the field of view otherwise the match with the stored pattern in the Inner Patch store will be degraded.

These then are the effects which set the performance of the tracker. Experience over several years indicates that if there is enough information in the picture for a man to recognise the target (usually a pre-requisite for an engagement) there is enough information to track it.

IMPLEMENTATION

Figure 6 shows a simplified block diagram of a typical Area Correlator tracker.

The video from the imager is sampled and digitised and then fed to the two Patch Control logic blocks. The sampling is defined by a master clock in the tracker control block. This clock is locked precisely to the synchronisation system of the imager to ensure that the same sample in each field represents the same position in the field of view.

The Inner Patch control selects the video representing the target as designated either by the operator's joystick – during initial designation, or from the latest measured target position when an update is triggered. This patch of video is then stored in the Inner Patch store.

The Outer Patch control selects the video centred on the target position for the Outer Patch store.

The correlator then accesses the data from both patch stores and carries out the correlation process. It then outputs the position of best match to the Target Position Logic and Tracker control. The Best Match Score is also passed to that block and to the Update Logic block.

The Update Logic block monitors the best match score and triggers the update process when necessary.

The Target Position Logic and Tracker Control block processes the target position information to refer it to the required datum, carries out the tracking filter calculations providing predicted target positions if necessary, and interfacing with operator controls and imager synchronising circuits. It also contains the master clock.

The circuit and logic are normally based on TTL integrated circuits and a micro-processor is used as a controller and to carry out the slower calculations. The correlator itself has to run at video rates and is normally constructed in hardware often with several parallel chains to reduce the delay times.

The shape and size of equipment depend to a large extent on the facilities required and the environment the tracker is to be used in. The photograph (Figure 7) shows a tracker, control unit and display monitor for a commercial environment. For an air-worthy equipment for a military application the tracker could be reduced to $\frac{1}{2}$ ATR short format.

MAN-MACHINE INTERFACE

The tracker removes the on-line tracking task from the operator. The requirements for control are for designation of the target, selection of tracking box size and switch to automatic track. Facilities are also available to adjust the tracking point and to adjust the box size during the automatic phase if required. This can be done without the machine losing lock.

The markers shown here may not be ideal and more work is required on the optimisation of both markers and joystick control laws for use in the cockpit of manned interdictor aircraft.

DEMONSTRATION FILM

A short film will be shown illustrating an EMI Electronics Ltd. tracker operating with a TV camera fitted to a two axis servo controlled mount designed and built by Evershed Power Optics Ltd. Chertsey, England.

CONCLUSION

The tracking system described based on the area correlator technique can provide a stable and accurate track of targets for use in airborne systems in conjunction with FLIR or TV imagers. It is of small size and can be built into standard format packages for installation in military aircraft.

The main advantages and disadvantages of this technique may be summarised as follows:-

Advantages : 1. It can handle targets with a wide range of characteristics

2. It can track background features to provide a ground reference

3. It does not require that the edges of the target are within the tracking patches. It can handle targets which completely fill the field of view.

4. It can adapt automatically to magnification and target aspect changes.

5. It provides highly tenacious tracking in low signal to noise conditions.

Disadvantages : 1. A very accurate designation is required if it is important to track a specific point on the target.

2. The target image must present a reasonable area within the field of view of the imager.

Applications for these systems include automatic control of laser illuminators for use with laser seeker semi-active weapons; control of imager sightline to provide stable pictures to aid recognition by locking onto the target or nearby background; and the heart of a passive homing missile system.
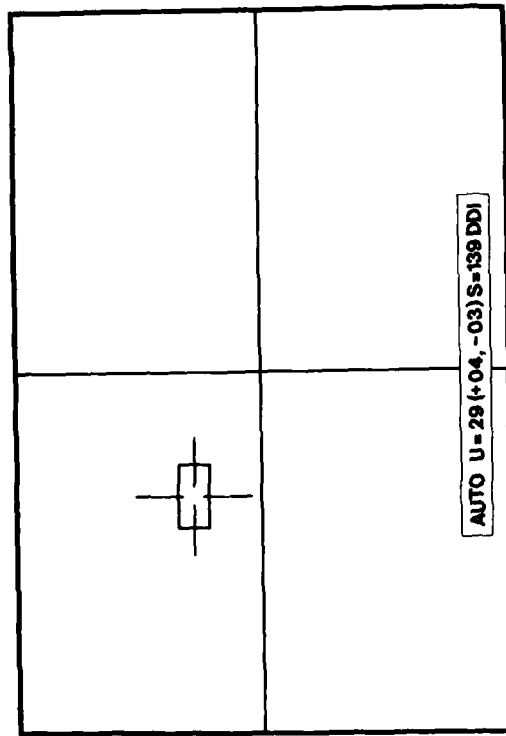
ACKNOWLEDGEMENT

AUTO U = 29 (+04, −03) S = 139 DDI

FIG.2 DISPLAY FORMAT

FIG.4 SEARCH PATTERN

LINE FORMAT

ODD FIELDS

EVEN FIELDS
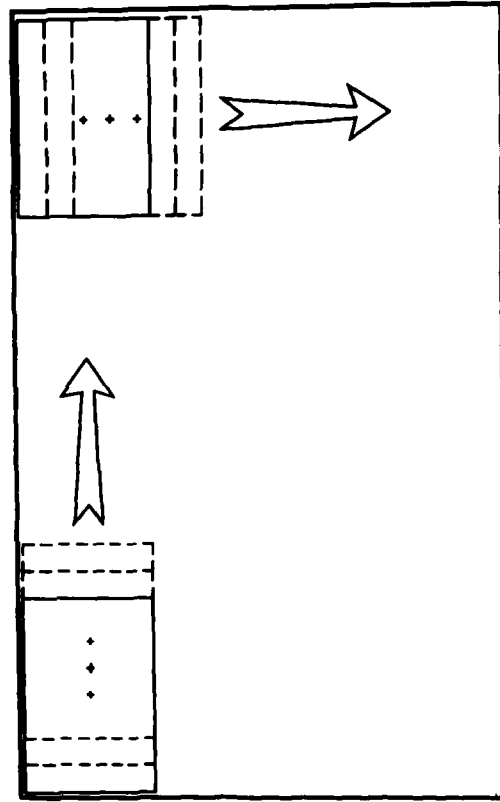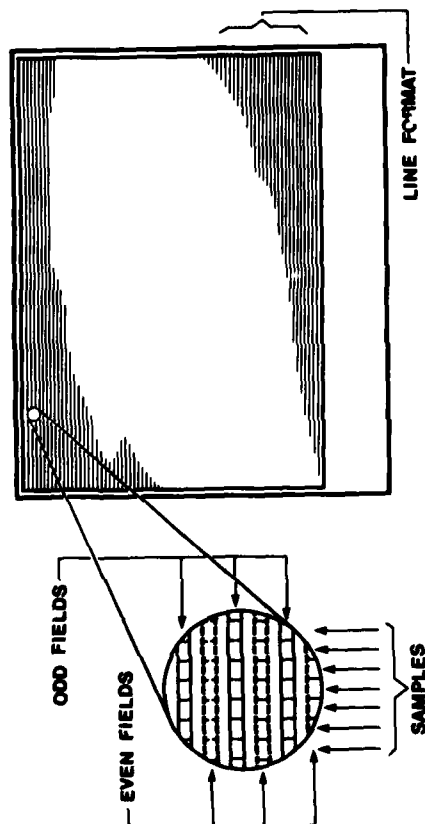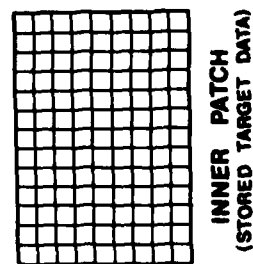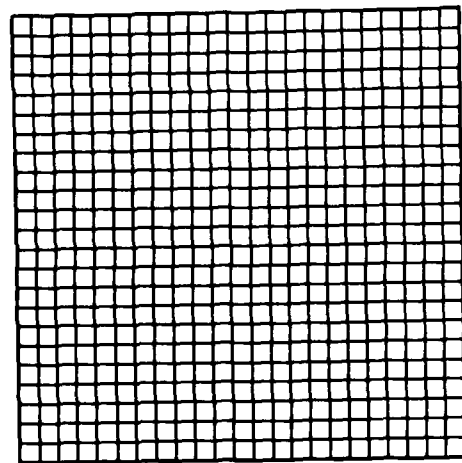
SAMPLES

FIG.1 VIDEO SAMPLING

OUTER PATCH
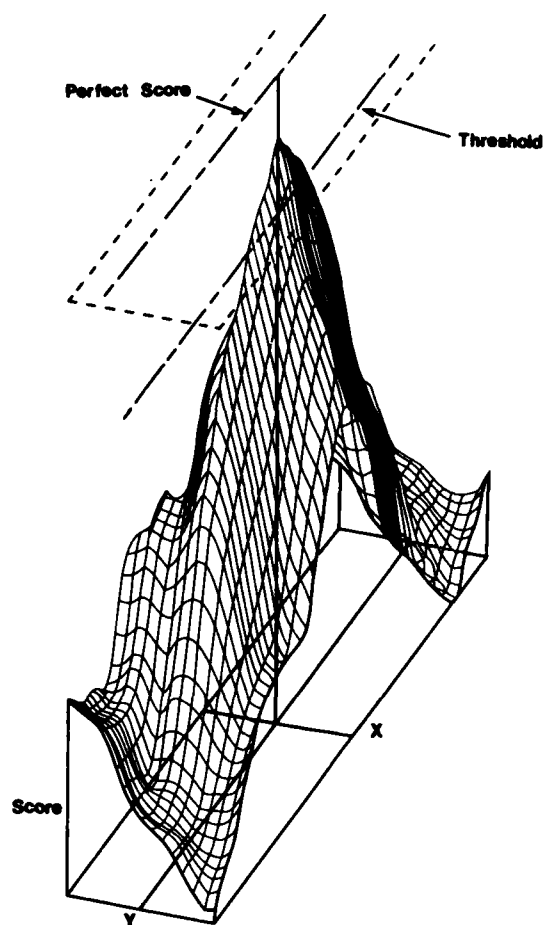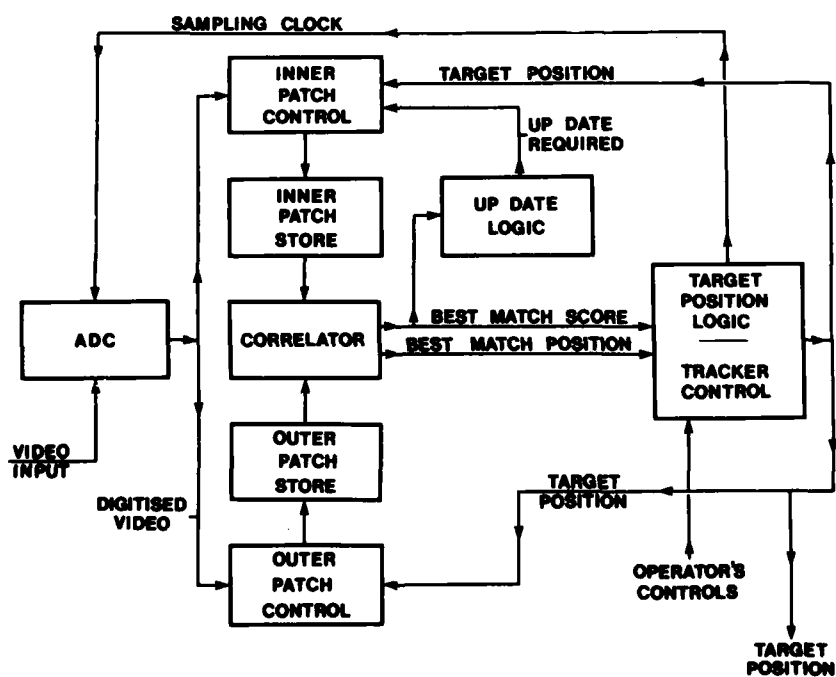(NEW FIELD DATA)

INNER PATCH
(STORED TARGET DATA)

FIG .3

FIG.5 CORRELATION SURFACES

FIG.6 SIMPLIFIED BLOCK DIAGRAM

27-8

FIG.7

A Real - Time Video Tracking System Based on Bayes' Rule

W. Wiemer
W. Laier
E. Keller

Messerschmitt Bölkow Blohm GmbH
MBB UA AE145 Postfach 801149
8000 München 80

Summary

Image processing techniques, quite generally, amount to a signal processing problem.
Target tracking is a special case of such a signal processing procedure. Essentially
based on the distinction between two processes (target and scenery), tracking is broken
down into three parts; signal acquisition and reduction, signal classification with
respect to target and scenery, and determination of the target coordinates with the aid
of the classified signals.

It will be shown, that the Bayes' decision rule works optimally in a statistical sense.
The target coordinates are determined by calculating the center of gravity using the
probability of presence of the object pixels. In addition the target coordinates are
filtered by means of a Kalmanfilter. This complete procedure is recursive and exhibits
an adaptive, learning behavior. Similar algorithms were used in other applications and
are known as unsupervised learning procedures.

Introduction

This paper serves to provide a brief summary on a special procedure used in image
processing, known as tracking.

Prior to describing this particular procedure, however, I should like to give some
general explanations on image processing.

As shown in Figure 1 the brightness (and possibly the color) measured, for instance,
in a square of 256 x 256 pixels, and the position in the measuring matrix are the only
data provided by the sensor. In addition one knows, however, and this is important for
further processing, that the terrain is static, that the object is compact and fluid,
i.e. the relationship between the pixels is not defined and that the target obeys a
moving model.

Hence, tracking can be broken down into two sections:
- signal acquisition and
- signal processing,
whereby additional data, which are not measurable but can be derived from physical laws
and are largely determined by the problem to be solved play a pivotal role in signal
processing.

But before we deal with the derivation of the theory the function of the tracker, based on Bayes' Theory, shall be demonstrated on a simple block diagram.

The principle functions of a tracker can be broken down into 4 blocks.
- signal acquisition and conditioning
- signal processing or classification
- determination of the target coordinates according to the classification results
- filtering and prediction of the hit coordinates based on the moving model of the target

These functions are summarized in the following simplified block diagram (Fig. 2), and can also be designated as a predicting algorithm for target coordinates.

As shown in figure 2 the algorithm is recursive. The a posteriori information obtained from the classificator is used as a new, a priori information. Moreover, Fig. 2 reveals that the signal acquisition is controlled by the classificator via the switching function u.

Remarkable analogies to control problems appear. For instance, the a priori and a posteriori information v and q, and the inputs to the classificator f_1, f_2 can be interpreted as state variables. Therefore the availability or estimation of definite initial conditions is necessary for initialization. This problem is well known in optimal filtering. Thanks to its optimization and based on its a priori information, the classificator correctly classifies and improves the measured signals with max. probability. Optimization is thus accomplished in a statistical sense.

The recurrence leads to the desired learning effect, i.e. the estimated position x, y tends to the correct value as the number of signals measured increases. This adaptive behavior is also characteristic of control problems. In contrast to most control procedures the algorithm used is however, extremly nonlinear and time-variant.

The brightness of the pixel constitutes the only input or feature in the tracker. The interrelation between the pixel positions are not taken into account, so that each pixel can be processed independently of the neighbouring pixel. This leads to a considerable simplification of the classificator.

Therefore, the processing can be easily extended to all pixels, i.e. background information can be included which accounts for the major part of the signals measured. For this purpose, it is imperative that the background is position independent and that it is mapped accordingly in the measuring matrix. This signifies that the position of the sensor must be fixed or stabilized according to a reference, as otherwise, the a priori information to the effect that the background is static can not be utilized.

Bayes Theory.

After these preliminary remarks we will now explain the Bayes' Theory. To solve the tracking problem, we have to decide whether a pixel belongs to the target or the scenery. Supposing that two competing processes exist the Bayes' Theory decides to which process a pixel belongs. And this is exactly what tracking is aimed at:
namely to determine whether this pixel belongs to the target or not.

The Bayes Theory solves this problem providing an optimal decision with respect to the cost function, whereby the mean cost risk is minimized and the costs are equally distributed to all possible error decisions. (Meyer-Brötz, Schürmann 1970)

This rule can be formulated as follows:

If
$$P_{\hat{j}} f_{\hat{j}} > P_j f_j \qquad j \neq \hat{j} \qquad\qquad (1)$$

is true, the process was measured.
This decision implies, that process $\hat{j}$ has been measured.

The definitions of this rule are as follows:

p_j = "a priori" probability that process j exists
f_j = the density of x, whereby x is the measured brightness of the process j
x = measured signal of a pixel and standardized brightness in the range of $0 \leq x \leq 1$
j = nested index for all possible processes $1 \leq j \leq n$
$\hat{j}$ = number of the process, which was selected according to the classification rule
n = number of possible processes

Equation 1 may also serve to calculate the a posteriori probabilities for measuring process i.

This probability to which the following formula applies

$$w_i \sim P_i f_i$$

is designated w_i.

In addition the equation applying to all probabilities

is
$$\sum_{i=1}^{n} w_i = 1$$

and therefore
$$w_i = \frac{P_i f_i}{\sum_{i=1}^{n} P_i f_i} \qquad\qquad (2)$$

For example the following decision is obtained by equation 1 for two processes, n=2

$$\hat{j} = 1 \quad \text{for } w_1 > w_2$$
$$\hat{j} = 2 \quad \text{for } w_1 < w_2$$

The calculation and meaning of p_i and w_i are unknown and shall be explained in the following:

Introduction of the terms 'visibility probability' and 'probability of presence.'

The density function of the terrain f_2, as a function of the position, as the terrain does not move and thus the dependence on the position contains important information. This does not apply to the density of the target f_1. It is calculated as a histogram which is position independent as the target is normally moving and changing its shape.

Representation of the density functions

In order to be able to calculate the density functions, the structure of this function must be assumed. A Gaussian distribution may be used for the density f_2 (terrain). The unknown parameters r and $\mathbf{6}$ must be estimated or learned.

$$f_2 = \frac{1}{6\sqrt{2\pi}}\, e^{-\frac{1}{2}\left(\frac{d}{6}\right)^2} \qquad d = x-r$$

The equations for r, $\mathbf{6}$ are as follows:

$$r(k+1) = r(k) + \alpha\,(x(k) - r(k))\cdot u \tag{4}$$

$$s(k+1) = s(k) + \beta\,(|x(k) - r(k)| - s(k))\cdot u \tag{5}$$

x = actual picture α, β, γ = gain < 1
r = reference picture
$\mathbf{6}$ = variance picture
k = time

Similar to equation 4/5 an estimation can also be set up for f_1. This is necessary, because f_1 may be subject to change.

$$f_1(k+1) = f_1(k) + \gamma(f_1(k) - f_1(k))\cdot u \tag{6}$$

u = 1   if the pixel is classified as a terrain point, for the time interval k
u = 0   if the pixel is classified as a target point, using time interval k.

Target coordinates.

As already mentioned, q consists of different probabilities, whereby one probability, referred to as probability of presence is used to calculate the target coordinates. Although there are several ways of determining these coordinates we have chosen the center of gravity of the probability of presence to represent the target point.

Based on the coordinates of all the pixels classified as target points, the center of gravity can be obtained by calculating the 0 and 1. order moments. To begin with, a target image $Z(i,j)$ is created, using the probability of presence q_1 and the choosable threshold value q_s.

$$Z(i,j) = \begin{cases} 0 & \text{for } q_1(i,j) \leq q_s \\ 1 & \text{for } q_1(i,j) > q_s \end{cases}$$

Using a track algorithm the process 'terrain' can be further broken down into foreground and background. An object may be present behind a foreground without being visible. A probability can be formulated for such hidden objects, and is referred to as probability of presence. This probability includes important information, since coordinates can be defined also for hidden objects.

However the foreground cannot be recognized unless the object disappears behind it, i.e., the foreground becomes visible or reconnoitred only, when the object is present. In this case only an insignificant a posteriori visibility probability (VP) w_1 can be calculated despite a high a priori VP (measuring probability p_1). The reason being, that brightness x does not belong to the object density f_1 but to the density of the scenery f_2.

Since an object cannot disappear from the scenery a certain probability can be formulated for the presence of the foreground.

To explain the mathematical derivations, the definitions determined for the single probabilities are as follows:

Definitions of the necessary probabilities.

p_0 a priori visibility probability for the foreground
p_1 a " " " for the object
p_2 a " " " for the terrain (fore- and background)

w_0 a posteriori visibility probability for the foreground
w_1 a " " " for the object
w_2 a " " " for the terrain

v_0 a priori probability of presence for the foreground
v_1 a " " " " for the object
v_2 a " " " " for the terrain

q_0 a posteriori probability of presence for the foreground
q_1 a " " " " for the object
q_2 a " " " " for the terrain

Although some of the values defined above are not required they have been indicated for the sake of completeness.

Based on these relations the following equations can be established:

$$P_1 = V_1(1-V_0) \tag{3a}$$

$$P_2 = 1-P_1 \tag{3b}$$

$$w_1 = P_1 f_1 / (P_1 f_1 + P_2 f_2) \tag{3c}$$

$$q_0 = V_0 f_2 / (P_1 f_1 + P_2 f_2) \tag{3d}$$

$$q_1 = w_1 + V_1 q_0 \tag{3e}$$

The next chapter demonstrates how f_1 and f_2 can be calculated using the measured brightness of the pixels.

Provided that v_0, v_1, f_1 and f_2 are known, all possible probabilities can be calculated.

Representation of the brightness as a density function.

As shown in the preceding paragraphs the density function f_1 (object/target) and f_2 (scenery) must be measured for classification purposes.

This can be achieved in several ways:
a)   the density function is position dependent
b)   the density function is position independent.

The density functions themselves, as a function of the position or not, can be determined parametrically or by counting (histogram). Three aspects play an essential role in determining the densities:

a)   data reduction
b)   exact description of the density
c)   simple hardware equipment

It is expedient to determine the sum of the rows and columns can then be calculated as follows:

$$ZS(i) = \sum_{j=1}^{jmax} Z(i,j)$$

$$SS(j) = \sum_{i=1}^{imax} Z(i,j)$$

imax, jmax height and width of the target window:

As a result, the zero and first order moments in i and j direction read as follows:

$$M_{0j} = M_{0i} = \sum_{i=1}^{imax} ZS(i) = \sum_{j=1}^{jmax} SS(j) \tag{8a}$$

$$M_{1j} = \sum_{j=1}^{jmax} j \cdot SS(j) \tag{8b}$$

$$M_{1i} = \sum_{i=1}^{imax} i \cdot ZS(i) \tag{8c}$$

And the center of gravity of the object is:

$$S_i = \frac{M_{1i}}{M_{0i}} \tag{9a}$$

$$S_j = \frac{M_{1j}}{M_{0j}} \tag{9b}$$

The filter shown in Fig. 1 is not only used to smooth the target coordinates, but also serves to predict the future position of the target by means of a moving model. These data which cannot be obtained by the Bayes' Theorem are required for the calculation of the a priori probability with the help of the a posteriori probability.

Calculation of new a priori probabilities on the basis of the previously calculated a posteriori probabilities.

Since learning procedures work in an adaptive way, the a posteriori information will react upon the a priori information of the following time step.

This leads to the following simple relationship:

$$V_0(i,j,k+1) = q_0(i,j,k) \tag{10}$$

  k: time
  i,j: position coordinates

Equation 10 presupposes that the process is not a function of the position and should therefore be applied with precaution.

As far as the foreground and thus f_2 and q_0 are concerned, no difficulties exist as the foreground is static and does not change in its shape.

This is different as far as the target is concerned, which moves and also changes its shape and size and, as a consequence calls for a prediction of the position and of the change in shape. This problem can be solved in several ways. Provided, that the target position can be predicted exactly with a Kalmanfilter and that the change in shape is slow, the following equation can be established:

$$v_1(i,j,k+1) = q_1(i-\Delta i, j-\Delta j, k) \tag{11}$$

whereby Δi and Δj constitute the difference between the predicted center of gravity $\hat{S}$ and the measured one $\tilde{S}$.

$$\Delta i = \hat{S}_i - \tilde{S}_i \tag{12}$$
$$\Delta j = \hat{S}_j - \tilde{S}_j$$

Realization.

To conclude with, some remarks shall be made with respect to the realization. An essential advantage of the Bayes' Theorem is, that the pixels can be classified independently of one another. The signals measured, can be classified immediately in the same order, as they are emitted by the sensor, without the necessity of a complete image.

Since the information of the image is evaluated in its completeness, and therefore both the probabilities of the target, and of the terrain are known, disturbances in the scenery are recognized and misclassifications minimized.

No realization problems occur if a clock of   5 MHz is used. The mathematical algorithm must be realized by means of pipeline techniques using a clock of 200 ns. A word length of 8 bits is sufficient and an automatic rescaling is necessary for the division in equation 2 only. The complete hardware is rather costly, as a large quick memory and expensive IC's are involved.

References

Günter Meyer-Brötz, Jürgen Schürmann, Methoden der automatischen
    Zeichenerkennung. R. Oldenburg Verlag München/Wien   1970

A.L. Gilbert, M.K. Giles, G.M. Flachs, R.B. Rogers,
Yee Hsun U, "A real-time video tracking system using image processing techniques", from
Processings of the 4th International Joint conference on Pattern
Recognition (Kyoto, Japan, 7.-10. Nov. 1978).

IMAGE PROCESSING = {
FEATURE ACQUISITION / SIGNAL ACQUISITION
FEATURE PROCESSING / SIGNAL PROCESSING
}
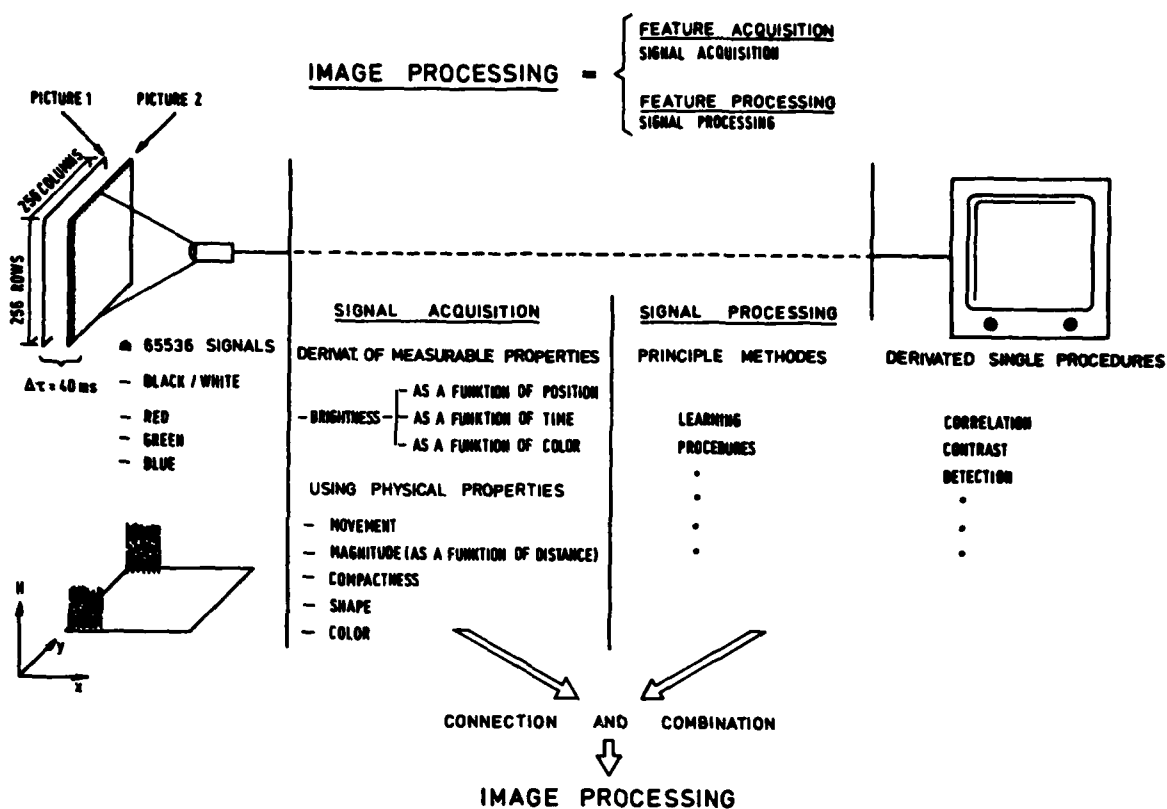
Figure 1

BLOCK DIAGRAM OF A TRACKER
BASED ON BAYES THEORY

$$V_0(i,j,k+1) = q_0(i,j,k)$$
$$V_1(i,j,k+1) = q_1(i-\Delta i, j-\Delta j, k)$$
$$i,j = \text{POSITION}$$
$$k = \text{TIME}$$

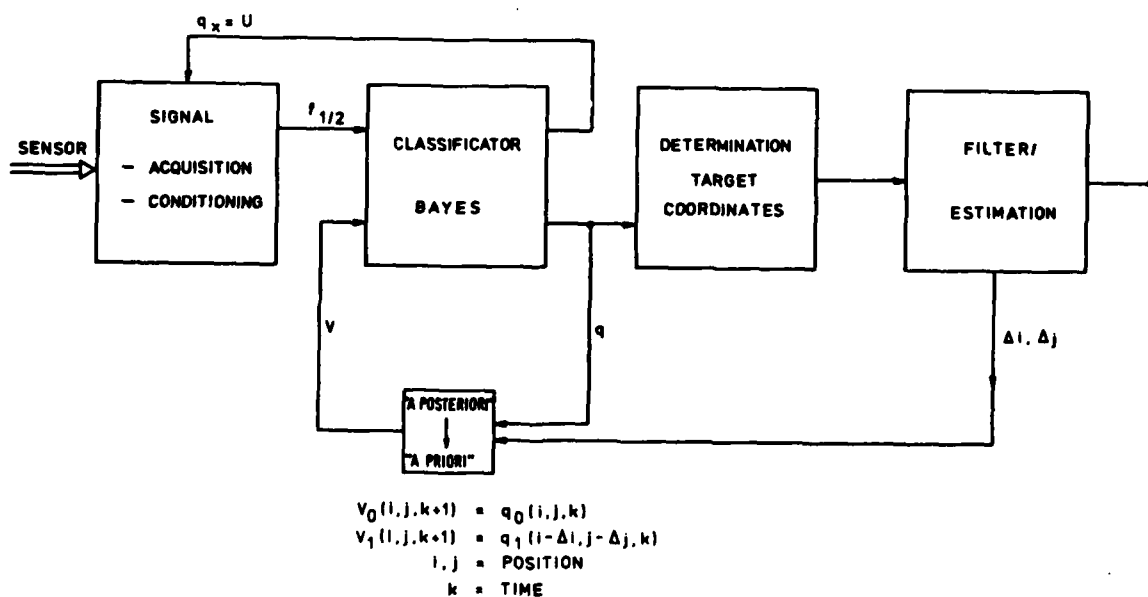Figure 2

Theoretical Presentation

Starting point: Optimal decision criterion

- Bayes Theorem

$$\boxed{p_j f_j \gtrless p_i f_i} \qquad \hat{j} \neq i$$

Calculation of the conditional probability:

$$w_i \sim p_i f_i$$
$$w_i = c \cdot p_i f_i$$
$$\sum w_i = 1$$
$$w_i = \frac{p_i f_i}{\sum p_i f_i}$$

$$\hat{j} = 1 \qquad w_1 > w_2$$
$$\hat{j} = 2 \qquad w_1 < w_2$$

Figure 3

PROBABILITY DEFINITIONS

PROBABILITY DEFINITION

P0	A PRIORI	VISIBILITY OF THE FOREGROUND	ESTMIATION NOT NECESSARY
P1	A PRIORI	VISIBILITY OF THE OBJECT	
P2	A PRIORI	VISIBILITY OF THE SCENERY	= 1-P1
V0	A PRIORI	PRESENCE OF THE FOREGROUND	ESTMIATION
V1	A PRIORI	PRESENCE OF THE OBJECT	ESTMIATION
V2	A PRIORI	PRESENCE OF THE SCENERY	= 1
W0	A POSTERIORI	VISIBILITY OF THE FOREGROUND	
W1	A POSTERIORI	VISIBILITY OF THE OBJECT	
W2	A POSTERIORI	VISIBILITY OF THE SCENERY	= 1-W1
Q0	A POSTERIORI	PRESENCE OF THE FOREGROUND	
Q1	A POSTERIORI	PRESENCE OF THE OBJECT	
Q2	A POSTERIORI	PRESENCE OF THE SCENERY	= 1

F1	DENSITY OF THE OBJECT	MEASUREMENT
F2	DENSITY OF THE SCENERY	MEASUREMENT

Figure 4

A posteriori probability calculation

A posteriori probability for the

— visibility of the object $\qquad W_1 = \dfrac{p_1 f_1}{p_1 f_1 + p_2 f_2}$

— visibility of the scenery $\qquad W_2 = 1 - W_1$

$\qquad\qquad\qquad\qquad\qquad\qquad p_2 = 1 - p_1$

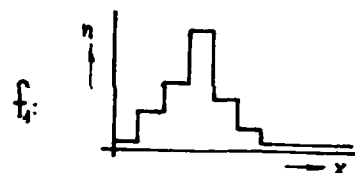— presence of the foreground $\qquad q_0 = \dfrac{V_0 f_2}{p_1 f_1 + p_2 f_2}$

— presence of the object $\qquad q_1 = W_1 + V_1 \cdot q_0$

Figure 5

Measurement, calculation, updating of the signals

Brightness density of the object: f_1
[assumption: position independent
$\qquad\qquad$ histogram]

f_1:

Brightness density of the scenery: f_2
[assumption: parametric representation]

$$f_2 = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{d}{\sigma}\right)^2}$$

Updating:

$$r(k+1) = r(k) + \alpha \left(x(k) - r(k) \right) \cdot u$$
$$s(k+1) = s(k) + \beta \left(|x(k) - r(k)| - s(k) \right) \cdot u \quad \Big]$$
$$\hat{f}(k+1) = \hat{f}(k) + \gamma \left(\widetilde{f}(k) - \hat{f}(k) \right) \cdot u$$

$d = |x - r|$

$\sigma = \sqrt{\frac{\pi}{2}}\, s$

Decision, classification:

$$u = 0 \qquad q_1 > 0.5$$
$$u = 1 \qquad q_1 < 0.5$$

Figure 6

TARGET TRACKING AND TARGET DETECTION IN

TV- AND FLIR-IMAGERY

M. Bohner
Forschungsinstitut für Informationsverarbeitung
und Mustererkennung (FIM/FGAN)
Breslauer Strasse 48, 7500 Karlsruhe
Fed. Rep. Germany

SUMMARY

This paper describes procedures for the detection of moving or stationary objects in single images and in image sequences, taken from a moving or stationary sensor. Primarily the objects are located on the ground in their natural environment so that simple detection procedures (e.g. detection by maximum intensity) cannot be applied.

For those situations a tracking system has been designed, simulated on a digital computer and tested with an equipment for real time application: At the beginning the object is detected by an human operator. From that moment it will be tracked automatically based on an evaluation of the correlation function between the actual scene and a memory which is continuously updated. The scale (distance between object and sensor) is determined separately and normalized by the location of object parts. The basic shortcomings of a correlation tracker system (e.g. image modification by foreground and background objects) have been eliminated in a system where in addition to the evaluation of the correlation function objects in the foreground or background are detected based on features like contrast, image differences, contour lines, shape and relative speed of the objects.

To substitute the operator at the beginning of the tracking an automatic detection and classification approach for the evaluation of single images has been simulated within a cooperative research project conducted by several NATO countries. For image areas obtained by a multiple threshold binarisation geometric features are computed and evaluated with a linear classification system showing detection rates of approximately 90% and classification rates of 50 - 60%.

1. INTRODUCTION

Object detection in single images and in image sequences is an important task for many applications (e.g. reconnaissance, battlefield surveillance, tracking of moving objects, homing systems). So far the object detection in those systems has been carried out by man (e.g. photo-interpreter, gunner). To reduce the time for the data interpretation or to substitute the man in dangerous situations automatic systems are needed. In this paper procedures for object detection in image sequences (correlation tracker) and in single images (multi-threshold object detection in thermal imagery) will be described.

2. CORRELATION TRACKER

Fig. 1 shows the basic components of a system for detecting and tracking objects in image sequences. The actual scene is continuously monitored by a sensor (e.g. TV-camera, IR-camera) and the sensor signals are processed on-line in an image processing unit where an operator can influence the processing by additional information about the object and the environment. The result of the processing is the position of a selected object in each image of the sequence and this position is used to control the sensor and its platform and to start system reactions due to the position of the object. The objects are stationary (e.g. tower, bridge, road) or moving (e.g. vehicle on the ground, aircraft, ship).

The platform with the sensor system can also be stationary (e.g. fixed on the ground or on a tower) or moving (e.g. mounted in a vehicle on the ground or in an aircraft). In the following paragraphs the image processing unit is described primarily for the tracking of objects on the ground because those situations are rather difficult to handle due to foreground and background problems. Detailed information is given in (BOHNER, M., 1976), (KAZMIERCZAK, H., 1978), (BACH, S., 1978).

2.1. Basic System

In the past many systems have been built, where the image processing was based on a simple evaluation of the intensity values of the object and its environment. After preprocessing of the sensor signals (e.g. windowing, weighting, linear or non-linear twodimensional transformation) the maximum intensity levels or other similar features like centre of gravity are used to detect the position of a selected object. Those systems work only successfully when a high contrast between object and background is given (e.g. aircraft in the sky, hot vehicles in a cold environment). Objects on the ground are often lost especially in the European landscape. Better results can be expected by correlation tracking.

Fig. 2 shows the principle mode of operation of a system based on the correlation function. An operator detects an interesting object (the car) in the moment $t(n)$, this object is stored in a memory as a reference and in the next image the correlation function between the actual scene at $t(n+1)$ and the reference memory is computed. The reference image has to be shifted within a search area which results in a field of correlation values $K(x,y)$.

These values are used to detect the object in the actual scene by a decision criterion and to update the reference memory for changing representations of the object.

Correlation function: As measure of similarity between the object o in the actual scene and the reference r the correlation function in its normalized version was selected.

$$K(x,y) = \frac{\Sigma(o_i - \bar{o})(r_i - \bar{r})}{\sqrt{\Sigma(o_i - \bar{o})^2 (r_i - \bar{r})^2}}$$

o_i, r_i picture elements of object and reference

$\bar{o}, \bar{r}$ mean values of o_i and r_i

The main benefit of this measure is its limited range of values (K = 1: r and o are equal, K = 0: r and o are statistically independent) and its independence to a linear transformation of the intensity levels (e.g. caused by changes in the illumination of the object).

System parameters: Parameters like number of grey-levels, resolution, size of reference memory, correlation and update rate have to be adapted very carefully to the special problem. As an example the influence of the memory size to the twodimensional correlation field is shown in Fig. 3. For a memory size too small for a special type of objects the autocorrelation has a sharp peak at the location of the object, but after a small variation of the object a detection is no longer possible (Fig. 3a). With the proper size of the memory the position of the object can be found in both cases (Fig. 3b).

Decision criterion: The position of the maximum value of the correlation function within the search area was selected as position of the object.

Update criterion: Two different criterions have been tested.
- The memory is updated as long as the maximum of the correlation exceeds a threshold, that means as long as a good similarity between reference and actual scene is given.

- The memory is updated when the maximum of the correlation is below a threshold, that means the memory is only updated when the appearance of the object has changed after a period of time.

The simulation showed that for such a simple correlation tracker the success of both criterions depends on the actual scene, no general preference of one criterion could be found.

Such a simple correlation tracker system has been simulated with a digital computer using sequences of real images in the visual range of the spectrum and has been developed with special hardware for real time application with a TV- or IR-camera as sensor.

2.2. Object Approach

When approaching an object the scale of the object representation is constantly increased. Because of the sensitivity of the correlation function to changes in scale, the object size in the image has to be reduced to an original value. The amount of this reduction can be determined by drift-references.

The drift-references represent parts of the object to be tracked and can be extracted from the first reference as conspicuous subareas (WINKLER, G., 1978) within the field of the total reference. In a special realization the original reference has been subdivided systematically into drift-references (Fig. 4a). After a period of time the location of each drift-reference is determined independently (Fig. 4c). From the location of all drift-references a value for the reduction of the actual image size can be derived. This system has been simulated successfully using test scenes as shown in Fig. 5.

2.3. System Modifications

The software simulation as well as the on-line hardware system revealed the basic problem of the method, demonstrated in Fig. 6 in principle. In the moment t(n) the reference memory contains the object without any foreground objects, whereas in the actual scene the car is partially hidden by the trunk of a tree. The maximum of the correlation function decreases but the object position can still be found with a high probability. If the update criterion is active in this moment, the trunk is learned into the memory as a part of the object. Therefore in the moment t(n+1) two peaks in the field of the correlation values are obtained (matches of car and trunk). Dependent on the size and contrast of the trunk a wrong object position may be found. A similar problem arises by big contrast changes in the background of the object. To overcome these problems several modifications have been simulated and evaluated.

In an environment with heavy contrast changes in the background a modified update procedure should be preferred. A new reference is composed of a part of the old reference and the actual image data. Because of the relative speed between a moving object and the stationary background, only the moving object is well represented, whereas background information is blurred. However with this update procedure new problems may arise when the representation of the object is changing very quickly.

In many situations where an object has been lost based on a detection by the absolute maximum of the correlation function a local maximum was found at the position of the object. Therefore the decision criterion was modified in such a way that all local maxima within the search area are extracted and weighted by a factor depending on the distance of the position predicted by an extrapolation of the object movement in the past. After the weight-

ing the original decision criterion can be used.

The prediction of the object movement can be controlled by a background correlation (Fig. 7). Image information in front of the moving object is stored in a background reference and this reference is correlated at a fixed position in the scene over a period of time. The correlation value at this position decreases when the moving object is crossing the background and increases approximately to the original value K = 1 when the object has passed. With the position x3 of the background reference and moment t3 of the correlation minimum the prediction of the object movement can be corrected.

In another modification the object reference is subdivided systematically into subreferences and the correlation values are computed for the total reference, a number of subreferences and some combinations of subreferences. The correlation values of the subreferences are very sensitive to small changes in the field of view of the object because of the small size of the corresponding references (Fig. 3). This feature can be used in a system demonstrated in Fig. 8 for segmentation into 6 subreferences and correlation with the total reference and subreferences 4, 5 and 6: The moving object is detected in each image of the sequence at the position of the maximum correlation value of the total reference. Small foreground objects however are detected by a rapid decrease of the correlation values of subreferences and in the corresponding areas the reference is not updated. Therefore the moving object can pass behind a foreground object without influence to the reference (Fig. 8b and c).

2.4.   Correlation and Object Detection

With the modifications of the simple correlation tracker some problems (sensor noise, foreground objects, contrast changes in the background) are less critical, but the basic origin of these problems is not done away. Moreover the simulation showed new problems which has come up with the modifications (e.g. quick changes of the object representation). To overcome this situation a new approach has been designed combining correlation and object detection methods. The object to be tracked is mainly located by the correlation of the actual scene and a reference. However the evaluation of the correlation values and especially the decision for the final object position is highly influenced by object detection methods where parts of the moving object and mainly foreground and background objects are detected and identified within a period of time. The detection and identification of those objects or parts of objects are based on features like

- intensity and contrast levels
- image differences of "adjacent" images within a sequence
- distribution of contour lines
- shape of special areas near the moving object
- relative speed between objects.

This approach has been simulated with a software system where the extracted foreground and background objects are taken into account by correlation and update templates (Fig. 9):

- Monitored by an update template foreground information is not learned into the reference memory

- Monitored by a correlation template the correlation values are only computed for picture elements which represent the object with a high probability, other picture elements are neglected.

The simulation with test scenes in the visual range (examples in Fig. 10) has been very successful, a real-time hardware system is under development.

3.   MULTI-THRESHOLD OBJECT DETECTION IN THERMAL IMAGERY

All approaches mentioned above need a man for the detection, classification and selection of the target to be tracked at the beginning of the sequence. Furthermore a target lost in a special situation within the sequence normally cannot be reacquired automatically and new targets entering the field of view in a later stage are not detected. To overcome these shortcomings objects have to be detected and classified automatically in single images.

Under Panel III of the NATO Defence Research Group the Research Study Group 9 (NATO AC/243 (Panel III)/RSG.9) is active in conducting cooperative research projects in image processing within the participating countries Canada, Denmark, France, Fed. Rep. Germany, The Netherlands, Norway and USA. Their first cooperative research project had the final goal of discriminating and classifying operational military targets in natural scenes from thermal imagery. Reaching this goal the above mentioned problems are solved.

In conducting this project the participating countries agreed on a common data base to which all algorithms developed in the different countries have been applied. This project has been terminated by a mutual evaluation of the various methods in respect to segmentation, detection and classification performance considering the requirements for real-time hardware systems. A final report is under preparation (SEVIGNY, L., 1980).

In the following paragraphs the German approach within this cooperative project is discussed in principle, a detailed description is given in (EBERT, A., 1980).

3.1.   Data base

The Alabama data base was selected as common data base. This data base consists of 43 images in the thermal range of the spectrum (3-5 μm and 8-14 μm) containing 84 vehicles as

> 40 tanks
> 29 armoured personnel carriers (APC)
> 15 jeeps

and other foreground and background objects. Typical examples of the imagery are shown in Fig. 11. The vehicles are always brighter than their local environment, but partially with a very low local contrast. Some vehicles are merged to a single blob with nearly no contrast to each other. The size of the targets (target area) varies from 10 to 500 picture elements with a grey-level resolution of approximately 8 bit in the total imagery.

For the cooperative research project the terms detection and classification has been defined in the following way:

> detection: discrimination between targets and background. The targets are not split into different types.

> classification: discrimination of detected targets into the three classes tanks, APC and jeep.

3.2.   Pre-Segmentation

The development of the segmentation algorithm was based on the idea that the moving targets are warmer or in the given representation brighter than the environment. In this case it is possible to separate the targets from the background by means of binarisation of the images using an appropriate threshold. Because of the big variety of the target intensities such a threshold has to be computed for each target separately. In Fig. 12 three thresholds have been used and it can be seen that for each target class a different threshold is optimal (a) for the tank, b) for the APC, c) for the jeep). There are well-known methods for finding such object adaptive thresholds if the contrast between the object to be separated and the background is fairly high and if some a priori information about the number and location of targets and their types is available. Such information was not taken into account and furthermore it was considered that in the future the contrast will become smaller due to better heat-camouflage techniques. Therefore the following procedure was chosen: A certain number of greylevels are selected systematically as thresholds and applied separately to the total image (level-slicing), so that a certain number of binary images is derived from each greylevel image. If the local contrast of a target is higher than the chosen distance between two adjacent thresholds, each target is at least separated from the background in one binary image. Feature extraction and object classification is applied to each binary image as described in the following paragraphs.

3.3.   Feature Extraction

As a result of the binarisation three types of regions are segmented:

> - those representing a target
> - those representing a part of a target
> - those representing background.

In order to be able to discriminate between these types and to classify the targets it is necessary to determine characteristic features for the different types of regions.

As the images are monospectral and as the application of textural features is not possible because of the size of a great part of the targets the only properties which have been used for the description of the binary regions (masks) are geometric features (size, shape and the course of the contour) and the distribution of the brightness within the masks. A list of all parameters with their mathematical definition is given in Table 1.

3.4.   Preclassification

After describing the obtained regions by the chosen features the desired target classes are found and separated by a classifier, background and foreground objects are rejected. After the selection of a classification system, the system parameters have to be optimized in a training phase. As agreed within NATO-RSG.9 for this training, 10 representatives for each class of targets had to be used. The targets were visually chosen in such a manner that the typical variations of each class are found within the training set. The binarisation of the selected representatives was done interactively with three different thresholds which were chosen so that

> - the contour is completely outside of the generated target region,
> - the contour coincide with the contour of the target region,
> - the contour is completely within the target region.

By this method of applying different thresholds for the same target two advantages are obtained:

> - the number of samples is increased which is necessary because of the 17 features which have been selected for describing the regions and

> - a target will be detected on more greylevels.

P_1	size of target region (SR)	
P_2	length of perimeter (LP)	
P_3	SR/LP	
P_4	max. radius of inertia (J1)	
P_5	min. radius of intertia (J2)	
P_6	J2/J1	
P_7	J1·J2/SR	
P_8	(J1 + J2)/LP	
P_9	angle α between the horizontal and the main axis	
P_{10}	mean greylevel within target region	
P_{11}	variance of greylevels within target region	
P_{12}	$(1/H) \Sigma \|W_{i+1}-W_i\|$	W_i width of line i H height of the comprehensive rectangle
P_{13}	$(1/W) \Sigma \|H_{i+1}-H_i\|$	H_i height of column i W width of the comprehensive rectangle
P_{14}	$(1/H) \Sigma \|WP_{i+1}-WP_i\|$	WP_i number of pixels in line i H see above
P_{15}	$(1/W) \Sigma \|HP_{i+1}-HP_i\|$	HP_i number of pixels in column i W see above
P_{16}	$(1/B) \Sigma \|X_{i+1}-2X_i+X_{i-1}\|$	X_i x-coordinate of pixel i on the perimeter B number of pixels on the perimeter
P_{17}	$(1/B) \Sigma \|Y_{i+1}-2Y_i+Y_{i-1}\|$	Y_i y-coordinate of pixel i on the perimeter B see above

Table 1: Definition of features

By means of the regression analysis a linear classifier was determined. This classifier $\underline{C}$ transforms the feature vector $\underline{x}$ into the discrimination vector $\underline{d}$:

$$\underline{d} = \underline{C} * \underline{x}$$

estimating the similarity of the region with the different classes. The similarity is op-
timal if the computed discrimination-coefficient is 1.0 and decreases with increasing
distance to 1.0. Moreover it must be claimed that the similarities with the other classes
are evidently smaller to obtain an unambiguous assignment to one class. A region is there-
fore accepted as a probable target of that class for which the highest similarity was found
with the additional conditions that the discrimination-coefficient is inside a fixed inter-
vall around 1.0 and exceeds the values of all other classes by a fixed amount.

3.5. Final Classification

An analysis of the Alabama data base revealed that the targets usually have sharp contrast
edges so that the binary target areas are stable for a certain number of adjacent thres-
holds. Background and foreground objects, however, may have a similar size and shape for
a single threshold but the pre-classification result changes rapidly with a modified bi-
narisation-threshold because of their smooth edge.

The investigations showed that for a selected stepwidth of 2 greylevels the false targets
usually are assigned to a class not more than two times. Therefore we accept a probable
target as a real target (detection) if it is assigned to any class of targets three or
more times. The class of the target (classification) is that one which was selected by
the preclassification for the majority of thresholds.

3.6. The Experimental Results

After the optimization of the algorithm its performance was tested with the complete image
data. A survey of the results of the different tests is given in Table 2.

In experiment 1 all features and the optimal bounds for the discrimination-coefficients
were used. Here it is possible to detect 78 vehicles out of 84 ones. That means that only
3 tanks, 1 APC and 2 jeeps were lost. On the other hand 3 false targets occurred. According
to our expectation the number of classified targets was evidently smaller. Only 25 tanks,
13 APCs and 12 jeeps were correctly recognized.

In experiment 2 the bounds for the discrimination-coefficients were modified in such a man-
ner that the requirements for targets were higher. Then only 36 tanks, 25 APCs and 13
jeeps could be located. But no false target was obtained. In addition the probability of
classification was equivalent to that of experiment 1.

In experiment 3 a reduced feature set and the optimal bounds for the discrimination-coefficients were applied. The number of detected targets was the same as in experiment 1. But only 17 tanks, 13 APCs and 12 jeeps could be correctly identified. In addition 13 false targets occurred.

In experiment 4 a further reduced subset of the feature set of experiment 3 and the optimal bounds for the discrimination-coefficients were used. Again the number of detected targets is equivalent to that of experiment 1. But the results of the recognition process were still worse than those of experiment 3. Only 8 tanks, 14 APCs and 9 jeeps could be assigned to the right class. Thus the number of classified targets dropped to 31. On the other hand the number of false targets was raising to 16.

classification experiment	0	1		2		3		4		5		6	
		D	C	D	C	D	C	D	C	D	C	D	C
Vehicles	84	78	50	74	49	78	41	77	31	78	49	70	46
Tank	40	37	25	36	23	38	17	38	8	38	21	34	22
APC	29	28	13	25	14	27	13	26	14	27	15	23	11
Jeep	15	13	12	13	12	13	11	13	9	13	13	13	13
false alarm		3		0		13		16		12		8	

0   total number of targets in image 1 to 43 of the Alabama data base

1   features P_1-P_{17}; threshold d_A

2   features P_1-P_{17}; threshold d_A^*

3   features $P_3, P_6, P_7, P_8, P_{12}, P_{13}, P_{16}, P_{17}$; threshold d_A

4   features $P_6, P_7, P_8, P_{16}, P_{17}$; threshold d_A

5   features $P_3, P_7, P_8, P_{16}, P_{17}$; threshold d_A

6   features $P_3, P_7, P_8, P_{16}, P_{17}$; threshold d_A^*

D   detection d_A = 0.5/1.1/0.2

C   classification d_A^* = 0.6/1.1/0.3

Table 2: Results for detection and classification

In experiment 5 another subset of the feature set as used in experiment 3 and the optimal bounds were applied. This time however not only the probability of detection but also the probability of classification was equivalent to those of experiment 1. The number of false targets was again evidently higher than in experiment 1.

In experiment 6 the same features as in experiment 5 and the modified bounds for the discrimination-coefficients applied in experiment 2 were used. Here only 70 targets namely 34 tanks, 23 APCs and 13 jeeps were detected while 8 false targets occurred. The result of the recognition process differs from that of experiment 5 only in the number of correctly classified APCs. Now 11 instead of 15 ones could be correctly identified.

The obtained results can be summarized as follows:

- approximately 90% of the targets of the Alabama data base can be detected with only a few number of false alarms

- 50-60% of the targets can be recognized

- the number of features can be reduced to 5 with only a small decrease in the detection and classification rate

- non detected targets are usually located in such a difficult environment, that there is nearly no chance for a detection by algorithms with a similar basic philosophy (this effect could be demonstrated by the mutual evaluation of the results within the cooperative research project).

REFERENCES

BACH, S.,   1978, "Detektion bewegter Objekte in zweidimensionalen Bildszenen"
GERLACH, H.,   FIM-Report No. 58
KAZMIERCZAK, H.,

BOHNER, M.,   1976, "Simulation eines Korrelationstrackers"
GERLACH, H.,   FIM-Report No. 44

EBERT, A.,   May 1980, "Discrimination and Classification of Operating
BOHNER, M.,   Military Targets in Natural Scenes from Thermal Imagery"
                    (Final Report - Fed. Rep. Germany), FIM-Report No. 77

KAZMIERCZAK, H.,   1978, "Automatische Zielverfolgung", FIM-Report No. 56

SEVIGNY, L.,   1980, "Discrimination and Classification of Operating Military
HVEDSTRUP JENSEN, G.,   Targets in Natural Scenes from Thermal Imagery", NATO-Report
BOHNER, M.,   (under preparation)
NAVARRO, F.,
OSTEVOLD, E.,
DEHNE, J.,

WINKLER, G.,   1978, "Measures for Conspicuousness", Computer Graphics and
VATTRODT, K.,   Image Processing Vol. 8, pp. 355-368
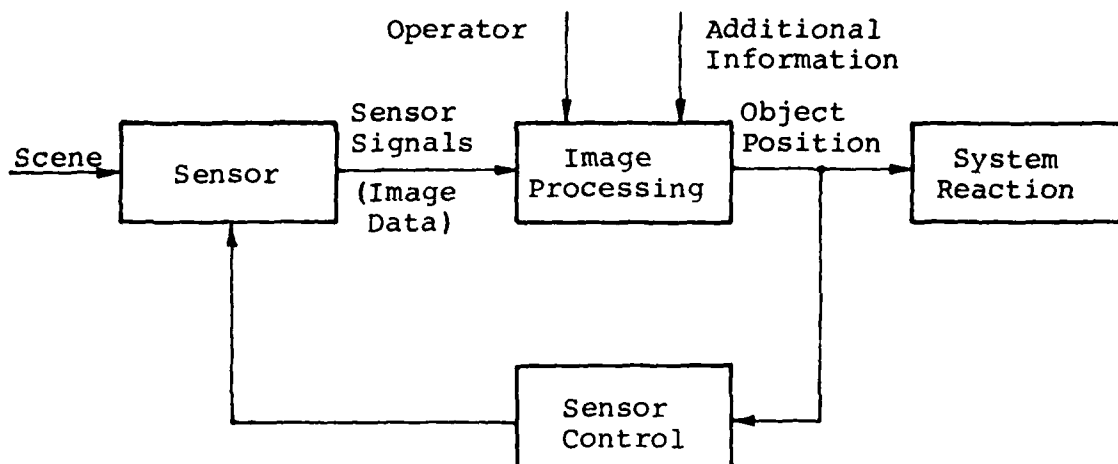
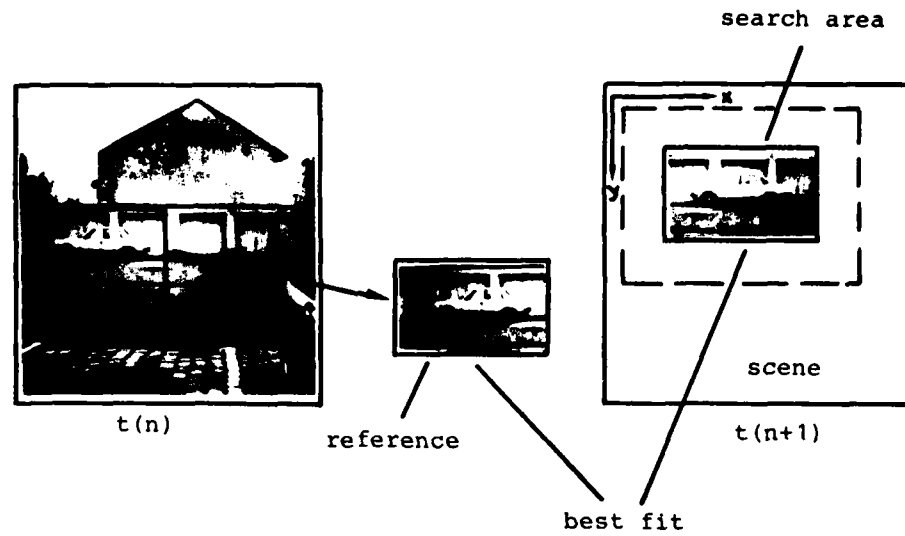Figure 1: Basic Components of a Tracker System
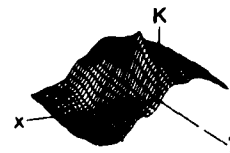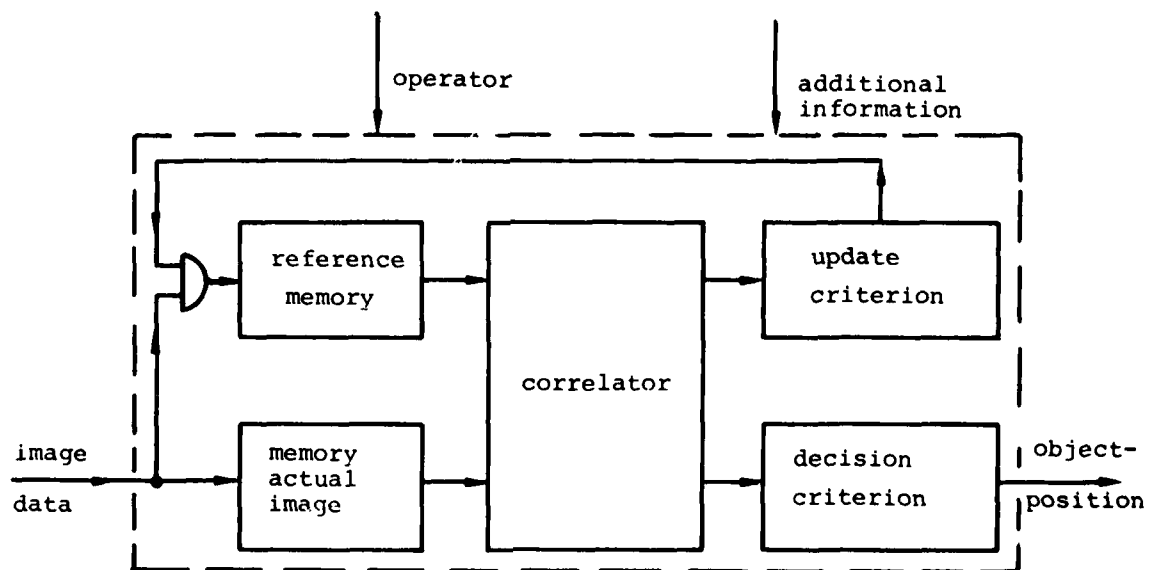
a) image sequence and reference

b) correlation field

c) blockdiagramm

Figure 2: Correlation-Tracker (Principle)

autocorrelation   crosscorrelation
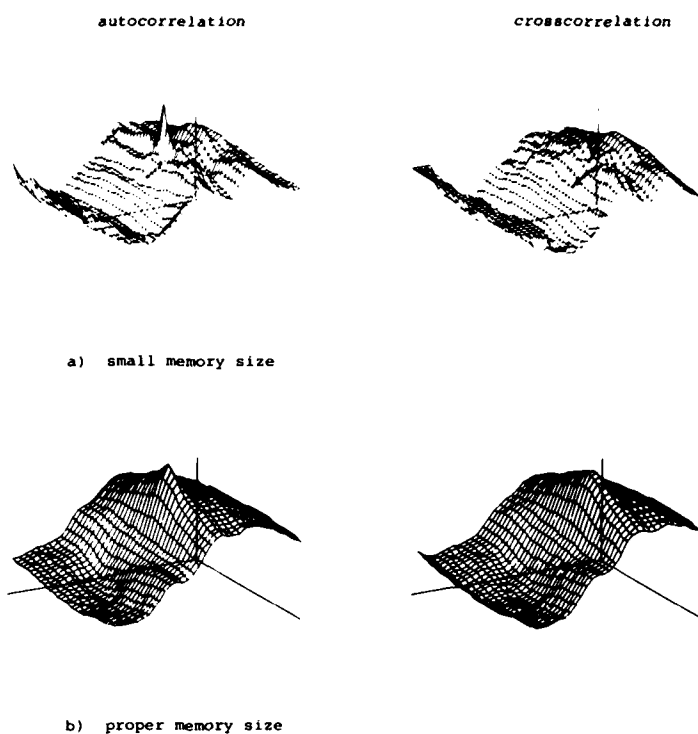
a) small memory size

b) proper memory size

<u>Figure 3:</u> Influence of Memory Size to Correlation Field

t(n)   t(n+1)   t(n+1)

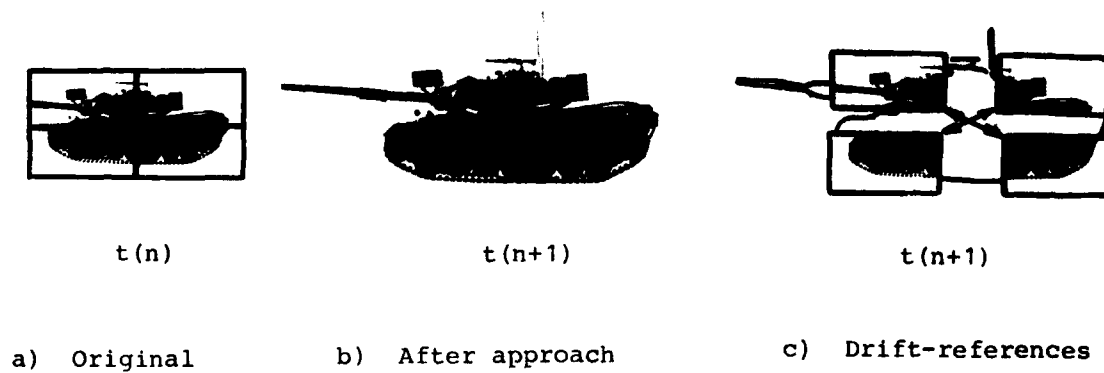a) Original   b) After approach   c) Drift-references

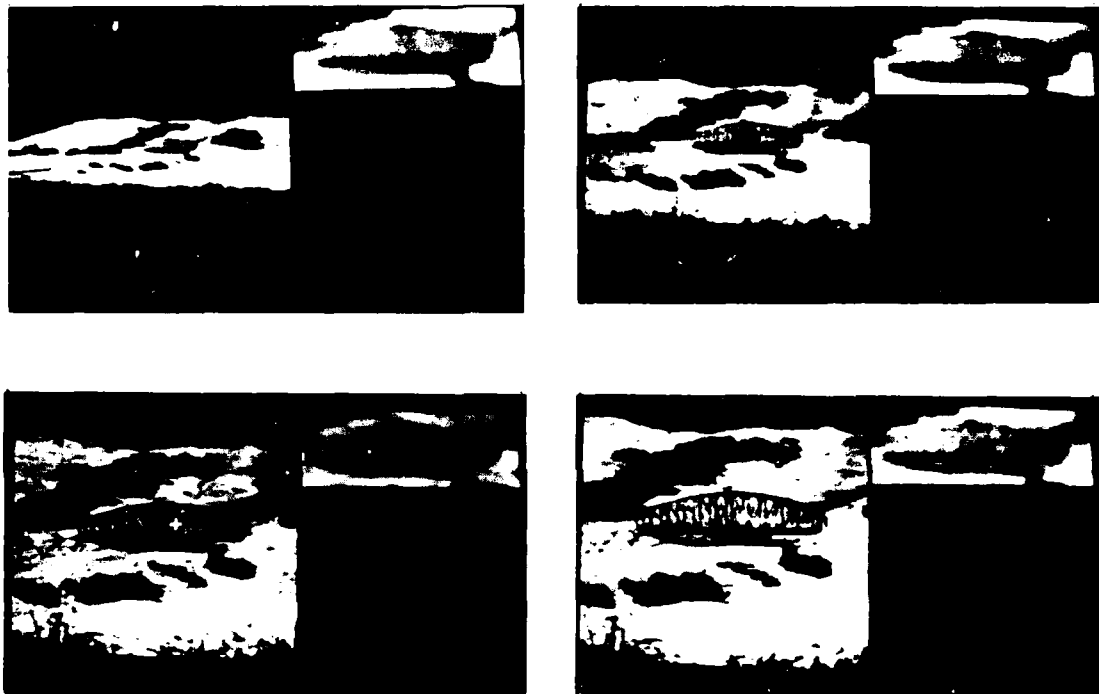<u>Figure 4:</u> Scale Control by Drift-References

Figure 5: 4 Moments in a Test-Scene for Object-Approach
(left: Actual Scene; right: Reference Enlarged)
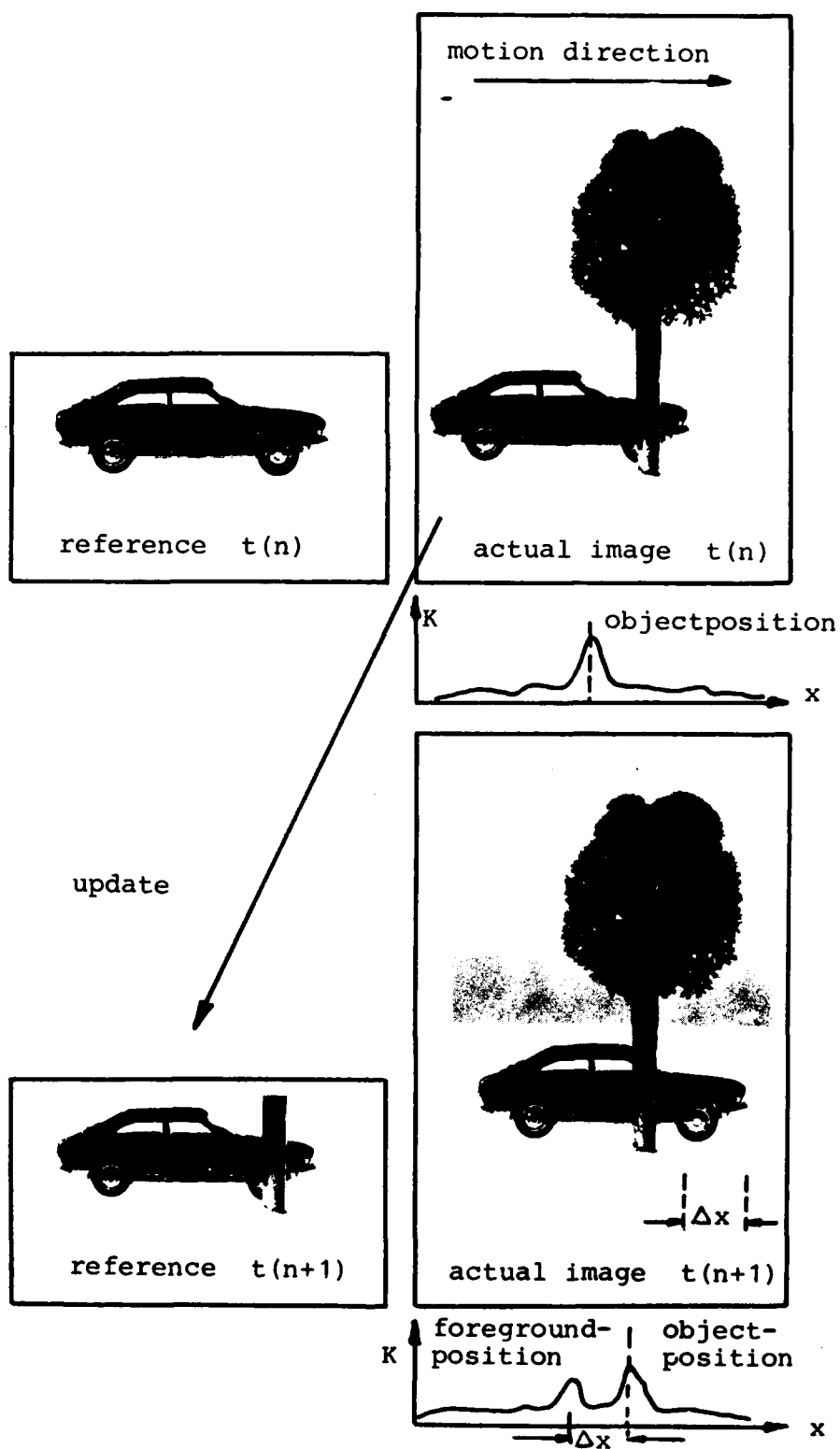
Figure 6: Influence of Foreground Objects
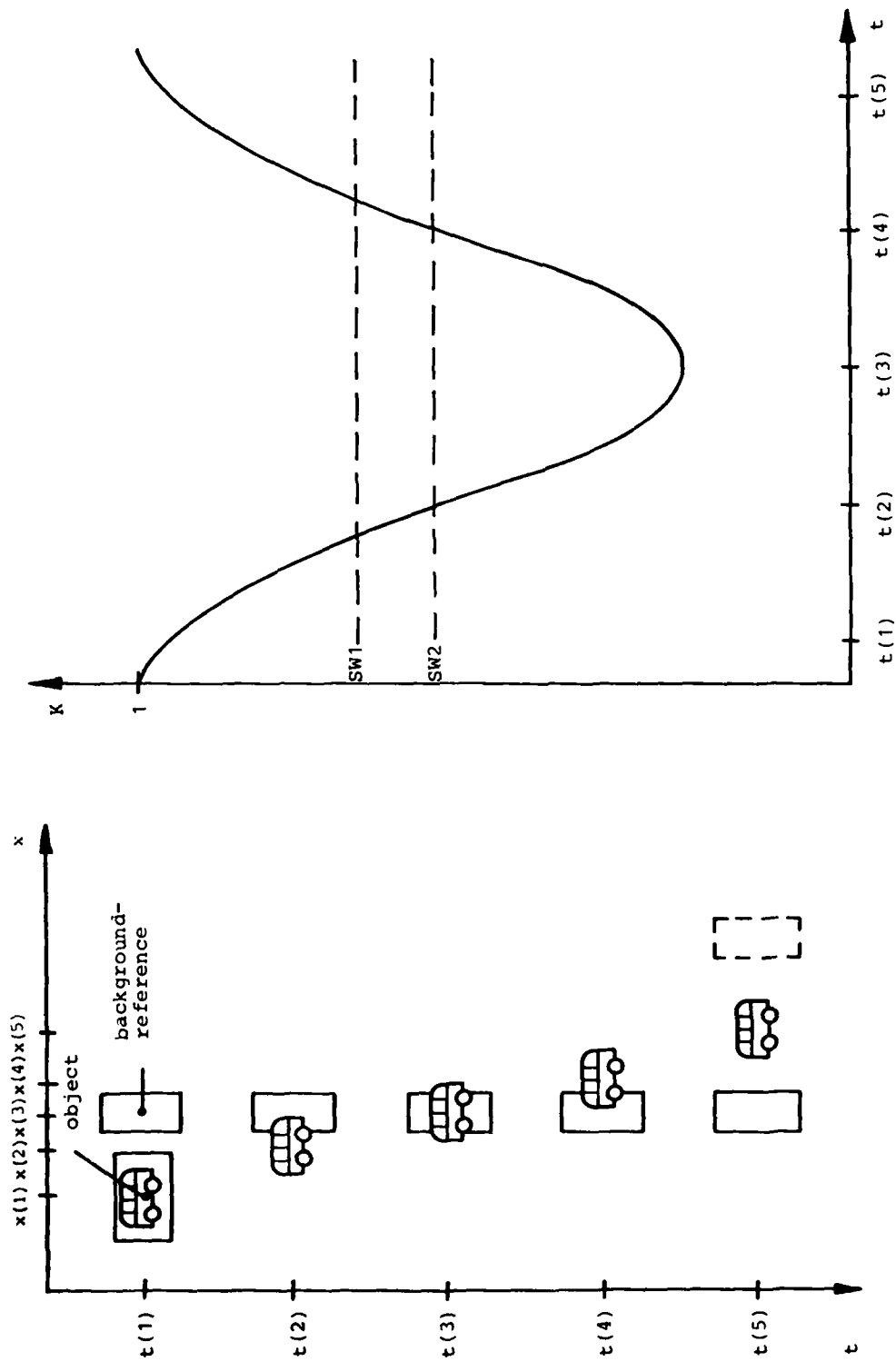
b) correlation between background
   reference and actual background

a) object movement and
   background reference

Figure 7: Principle of Background Correlation

a) Reference segmentation

foreground object

NU = No Update

b) Update procedure

total reference
subreference 4
subreference 5
subreference 6
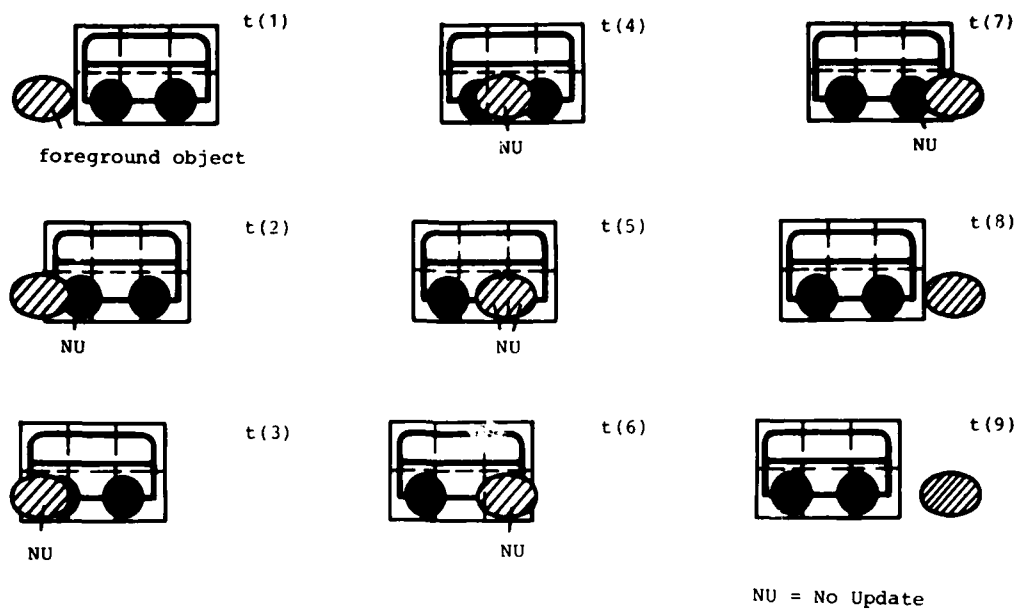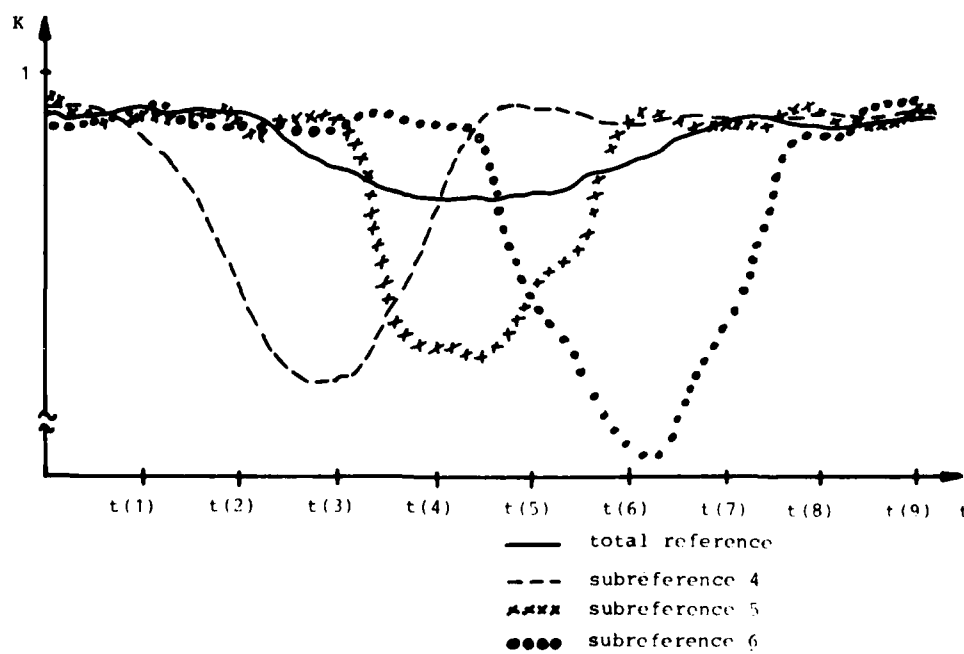
c) Correlation values for total reference and subreferences

Figure 8: Subreference Correlation

Figure 9: Template Correlation

Figure 10: Examples for Test-Scenes

Jeep, APC, Tank
(low contrast)

Tank
(high contrast)

Tank, APC
Jeep (very low contrast)

Tank
(large size)

Tank, APC
(small size)

Jeep, APC, Tank
(merged to one blob)

Figure 11: Examples of the Data Base

Figure 12: Binary Images of Image 1 in Fig. 11 with 3 Different Thresholds Applied to the Total Image

TARGET IDENTIFICATION
OF AIRCRAFT USING
IR/TV-SENSOR-IMAGES

Dipl.-Ing. Dirk Meyer, Dipl.-Math. Manfred Müller, Dr. Andreas Weimann

ESG Elektronik-System-Gesellschaft mbH
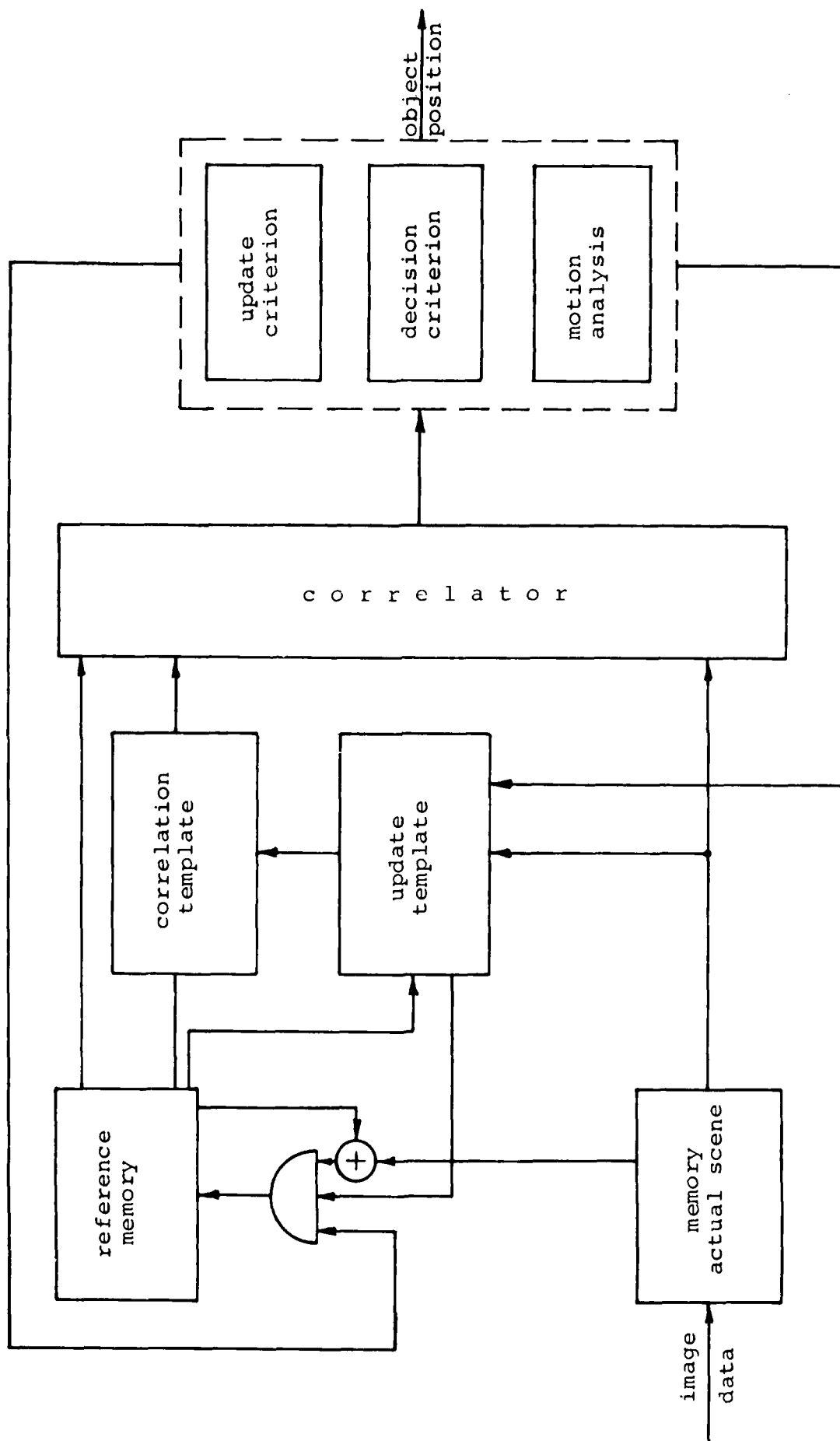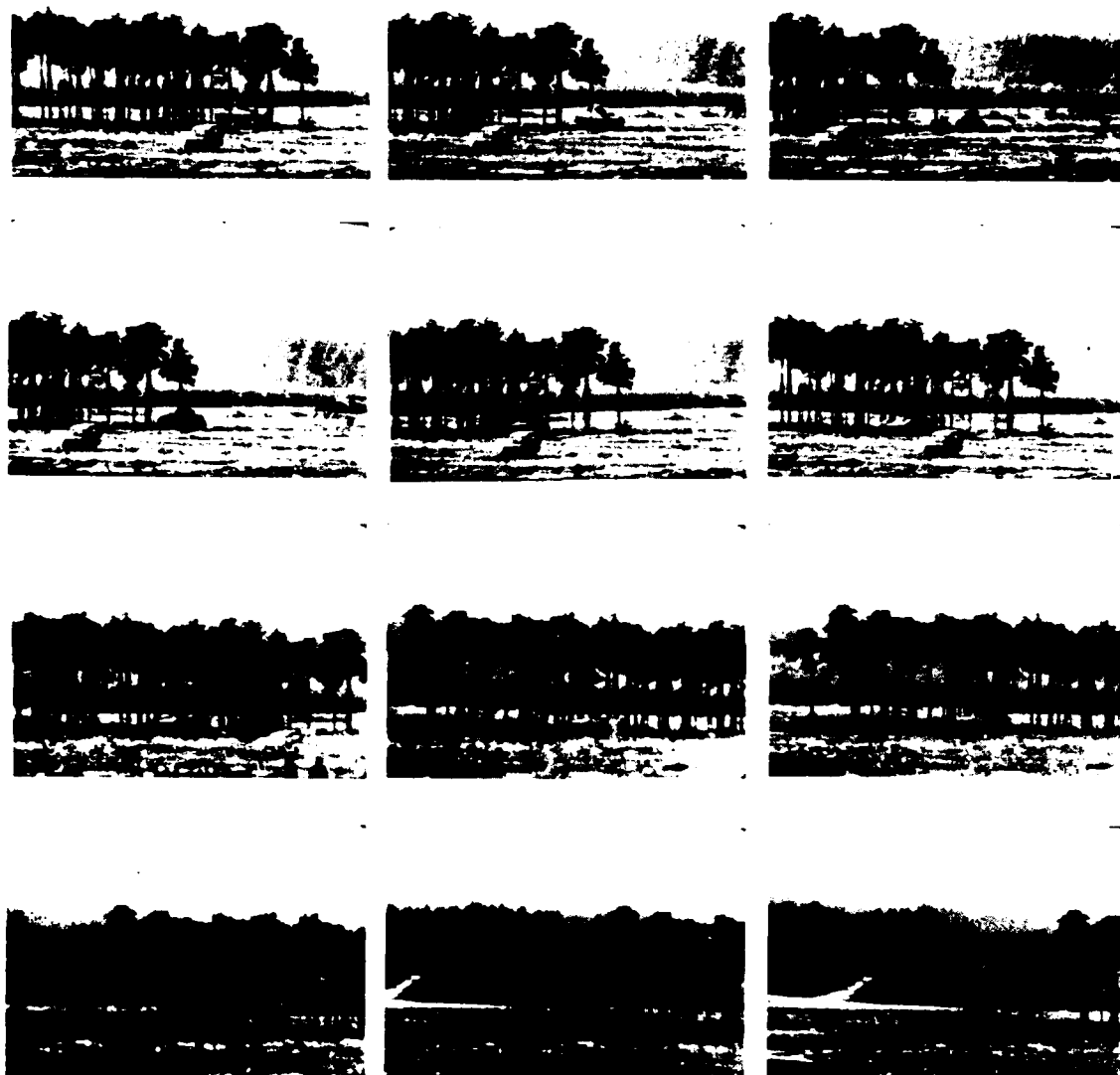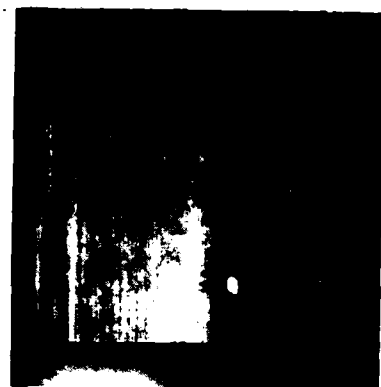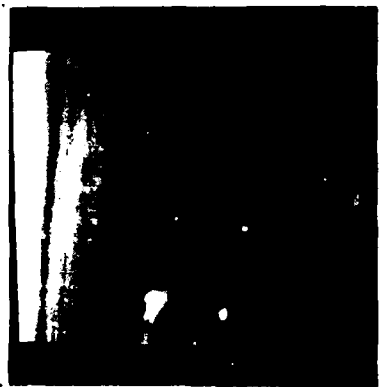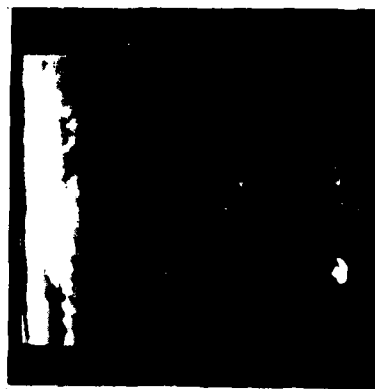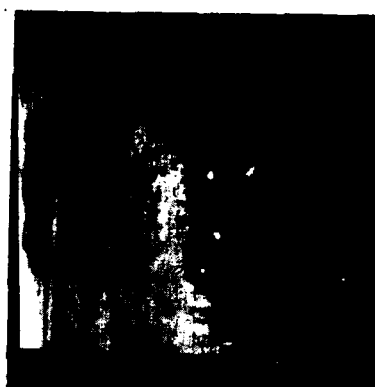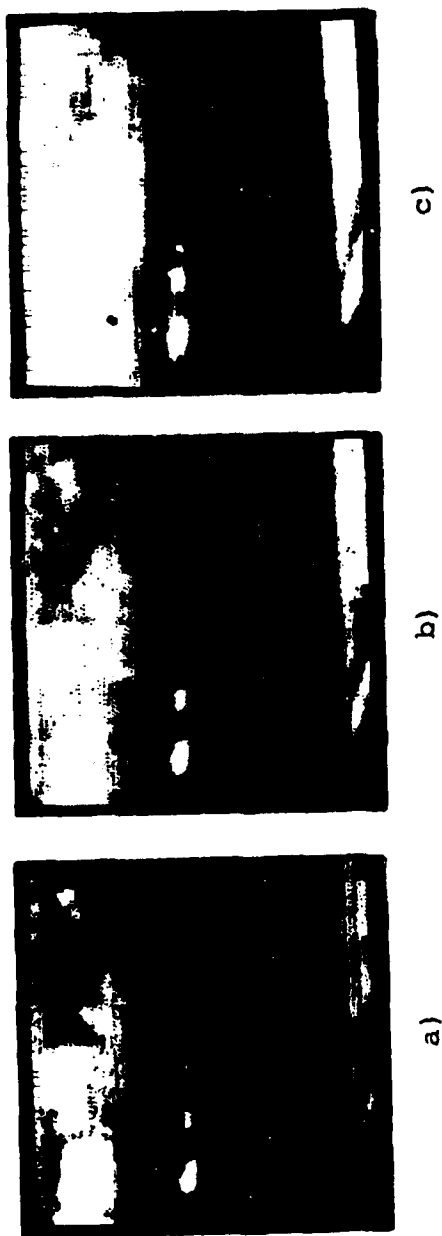Vogelweideplatz 9, 8000 München 80, W.-Germany

SUMMARY

In complex air combats multiple target tracking and target identification is a very dif-
ficult task for a pilot. To support the pilot, procedures for automatic tracking and
identifying processing multi-sensor-data and sensor-images should be developed.

The paper presents a procedure for non-cooperative target identification of aircrafts by
pattern recognition applied to TV/IR-sensor-images. By a multi-sensor-tracking of high
accuracy the position of an aircraft is very well known. An imaging-sensor with a narrow
field of view is pointed to the aircraft if it is in range. The image is then processed
for significant parameters. The contours are extracted by adequate methods of image pro-
cessing and evaluated for geometric relations which are independent of the projection
angle and specific for individual types of aircrafts. The application to typical air-
craft contours shows that the ensemble of the extracted parameters can identify different
types of aircrafts with some significance.

Furthermore an algorithm based on momentum invariants is proposed. The algorithm was
implemented and tested on a digital computer by means of simulated noisy images. Some
examples of identification results are presented and discussed.

1. INTRODUCTION

Future airborne weapon systems will have a higher degree of automation and integration
for target detection and identification. During air combat, especially dog fight
situations, the pilot's workload is extremely high. He has to perform concurrently a
lot of tasks under severe physical conditions in a high-g-maneuver environment. Under
these conditions it is necessary to support the pilot in target indentification by
automatic and/or semiautomatic procedures.

An autonomous air combat is characterized by the fact, that only a primary radar can
detect and localize a target within the range of an air-to-air missile. But the reso-
lution of a radar does not allow to distinguish between friend and enemy types of
aircrafts. A secondary radar (IFF) only can recognize if the target is cooperative
(answering) or non-cooperative, but is does not make sure that a non-cooperative tar-
get is an enemy. On the other hand the resolution of the direction of an secondary
radar is very low. Therefore it cannot discern targets which are close together. In
combat situations with friend and enemy aircrafts it is still today often necessary
to close up for a visual target identification /1/. By electro-optical means one
tries to have larger range for target identification.

The objective of this paper is to propose some methods for multisensor target detec-
tion and target tracking and performing target identification by pattern recognition
applied to electro-optical sensor images.

2. MULTI-SENSOR-APPROACH

Information about an airborne target can be collected by different sensors, for
example

- radar
- radar warning
- TV/IR-imaging sensor

As shown in figure 1 every sensor gets information about the target, which is cor-
rupted by noise. For evaluation of this information the signals of an individual
sensor may be filtered. But inspite of displaying these filtered signals directly to
the observer, who only can obey the outputs of one information channel at a time, it
is more convenient to evaluate the filtered information of all sensors together in a
sensor fusion by automatic correlation/combination of the information of all sensors
and known information about the objects to be identified stored in a catalog. Evalu-
ating the total information the correlation process has higher probability of iden-
tifying the target. The results then are displayed to the observer who probably has
to perform additonal interactive functions concerning his greater abilities in
interpreting complex situations.

Applying this concept of sensor fusion it is obvious that the characteristics of the
sensors complete each other. Radar and radar warning can detect a target in far
ranges, but they cannot positively identify a target on a non-cooperative basis be-
cause the target appears usually only in one resolution cell (this might not be true
for jet engine modulation and for a more complex radar with highest resolution). If
the target is closing up the radar can give information to the imaging sensor to
direct them onto the target.

If the target is in the range of radar and TV/IR-sensors the information of these sensors can be used to accurately track the target. For multiple target tracking the characteristics of the sensors complete each other. The radar has a high resolution in range and range rate but a low update rate when operating in a scan mode.

The TV/IR-sensors have a high resolution in azimuth and elevation and a extremely high update rate. Combining this information by applying an appropriate tracking filter one can accurately track the targets. From the target track the kinematics of the flight path can be concluded and the characteristics of the target derived such as:

- velocity
- flight hight
- g-maneuver
- climb rate

These from the flight path extracted data can be compared with stored data of target types, i.e. that an aircraft type of the catalog does not fulfill the performance of the target and therefore it cannot be that type.

If the target is even closing up the TV/IR-sensors should be switched to a narrow field of view and by knowing the target position very accurately directed onto the target. If the target is represented by an appropriate number of pixels in the image it can be evaluated by pattern recognition which is discussed in some approaches in the 4th chapter. To support the digital image processing the accurately known position of the target can be used to reduce the amount of computing by processing only that part of the image in which the target is located. Furthermore the known flight path of the target together with the known own flight angles can be used to roughly estimate the aspect angle of the target which has advantage for the pattern recognition.

3. SENSOR CHARACTERISTICS

Airborne target identification can be performed by the following types of passive imaging sensors /11/:

o TV-systems (Daylight-TV, LLL-TV)
o Thermal imaging devices (FLIR)

The quality of imaging systems can be described by means of the following characteristics/parameters:

- SNR (signal to noise ration)
- geometrical resolution
- minimal contrast at display (for TV),
  MR ΔT (minimal resolvable temperature difference)
  (for imaging IR, FLIR)
- dynamic range (detectors, display)
- saturation effects
- lag, image blurring
- scan format (number of lines, frame rate interlacing)
- video output (seriell, parallel)
- degradation by atmospheric effects
- night capability (bad weather capability)

PERCEIVABLE IMAGE CONTRAST

In order to perceive a certain apparent image contrast C_i this contrast (or the modulation of the signal carrier $\bar{n}$, e.g. signal current, image intensity etc.) must be superior by the factor k_r to the SNR:

$$C_i \geq k_r \frac{\mathfrak{G}(n)}{\bar{n}}$$

with $\bar{n}$ = mean value of signal carrier,
$\mathfrak{G}(n)$ = variance of n (fluctuation, noise)

This relation is valid for each type of perception, even for film granulation, and image (background) structure /3/.

The image quality is not only degraded by noise, it is also affected by the transfer characteristics of the imaging system (mostly low pass characteristics). This influence can be described by the MTF (modulation transfer function). In respect to a characteristic parameter (e.g. spatial frequency, angle resolution) the transition from the object contrast C_o to the apparent image contrast C_i can be described as

$$C_i (\alpha) = C_o \cdot T_m (\alpha)$$

$$T_m (\alpha) = \text{MTF curve}$$

For the recognition of an object (with means of a display) must be fulfilled:

$$C_o = \frac{C_{i\,min}(\alpha)}{T_m(\alpha)} \geq \frac{k_r}{T_m(\alpha)} \cdot \frac{\mathfrak{G}(n)}{\bar{n}(\alpha)} / \text{out}$$

This equation is to be normalized by means of the noise figure F,

$$F = \frac{(\;\sigma(n)/\bar{n})^2\;/out}{(\;\sigma(n)/\bar{n})^2/in} = \frac{1}{\eta_N}$$

with η_N = quantum yield

$$c_o \geq \frac{k_r}{T_m(\alpha)} \cdot \frac{\sqrt{F}}{\dfrac{\bar{n}(\alpha)}{\sigma(n)}/\;in}$$

In this equation the noise figure F is to be used corresponding to the type of image tube or detector. This noise figure is as that equation demonstrates, a significant factor to characterize the system performance. However, additionally the system is to be described as a function of resolution and object structures, most useful by the MTF.

A recent experimental investigation of the recognizability of details in TV-pictures which are impaired by noise confirms, that the SNR should be superior to 10 dB....20 dB- depending on the details to be recognized. /6/.

SOURCES OF NOISE

The generation of noise is based on two effects:
1. thermical noise and 2.statistically fluctuations (quantum noise). Thermical noise is generated by the load resistance of the detector and by the preamplifier.

The quantum noise is caused by two sources; one internal source (that is the detector dark current) and one external source (that is the inherent quantum noise of the received radiance).

In every imaging system the detector current consists of three parts:

$$I_D = I_s + I_b + I_d$$

- signal current: I_s.
  This current is generated by irradiance power carrying the desired information
- background current: I_s.
  This current represents the undesired radiance from the background
- dark current: I_d.
  This current is generated internally by the detector.

MTF (modulation transfer function)

The total MTF of an operating system results from a convolution of the MTF of the atmosphere with the MTF of the EO-Sensor (both are of low-pass-characteristics). The MTF of the sensor is determined by the MTF of the optics, by the MTF of the detector (e.g., detector array or image tube), and by the MTF of the video-amplifier. The total MTF can be slightly adjusted by controlling the frequency characteristic of the video amplifier.

Usually optics are diffraction limited and the detector characteristics are of most influence. Figure 2 shows the modulation depth of a 2/3"- and of a 1"-Vidicon in respect to the line-number.

RESOLUTION

The max. resolution of an EO-system is to be calculated from the necessarily SNR and the MTF-curve.

For TV-image tubes the SNR degrades with decreasing illuminance on the face plate and with decreasing image contrast, and, therefore, the resolution degrades too, as shown in figure 3. These or similar curves are shown in the data sheets of the manufacturers. (The IR-Sensors are of similar behavior).

The highest resolution, that can be obtained at a certain contrast level determines the recognizability of objects and their detail structures. This relation is valid for visual observation via monitor and for automatic image processing.

High stabilization quality is necessary to avoid performance degradation due to vibrations of aircraft and shock.

OPERATIONAL CHARACTERISTICS

For a realistic application the optical aperture is limited by technical and financial reasons. For diffraction limited optics the maximal angle resolution is given in figure 4. The corresponding ranges for imaging sensors with these optics are displayed in figure 5 for several target sizes, but without the atmospheric influences. Under these conditions we obtain with a high resolution TV-System ($\alpha \approx 0,025$ m rad) identification ranges from approximately 4 km at head-on aspect up to more than 15 km for full target size.

But under various conditions the sensor range can be limited severeously by the extinc-

tion of the atmosphere. The atmospheric extinction is very high for horizontal path near ground, but rapidly decreasing with height.

Therefore the atmospheric extinction in air combat situations is - despite of clouds - mostly extremely low and than the ranges are only determined by the sensor performance. For a model clear standard atmosphere contours of constant atmospheric transmittance are calculated /5/ at different radiation wave lengths (figure 6). The sensor range is severeously limited for horizontal path near ground, but increases rapidly for slant range with increasing height of the observer.

The sensor range at high atmospheric transmittance is only limited by the geometrical resolution and the SNR. (For special purposes the transmittance can be calculated by using the LOWTRAN-Program.

Figure 7 /4/ displays the probability of target detection in respect to the target distance (horizontal path on sea level) or a TV-camera (625 lines, 3/4"-Vidicon (resolution $\approx 0,06$ m rad)): The dashed line represents the limit given be the SNR, the dotted line displays the ranges, for 5% contrast at the entrance of the optics caused by atmospheric extinction. For a target with high contrast (K = 1; m = 100%) the resulting range (curve 1) is also affected by the atmospheric extinction and by the SNR of the Sensor. For a target with low contrast (K = 0,3; m = 22%) the possible range is only determined by the atmospheric transmittance.

For application of IR-Sensors - despite of the better atmospheric transmittance - the principals are very similar as figure 8 displays. Curve 1 shows the limitation by the SNR of the detector, curve 2 is given by the geometrical resolution; the resulting curve 3 is obtained by the convolution of 1 and 2.

4. SOME CONCEPTS OF PATTERN RECOGNITION FOR TARGET IDENTIFICATION

In the discussion of the multi-sensor-approach for target detection and localization the importance of electro-optical imaging sensor for identification was pointed out. The previous chapter dealt with the characteristics of such sensors. In the concept to be discussed it is supposed that a single static image is available in which the target aircraft is represented in a sufficient number of pixels with sufficient contrast to the background. That image is digitized and used for target identification by digital image processing.

Three basic concepts of pattern recognition for identifying aircrafts will be discussed in some detail:

- first concept using invariant length ratios
- second concept using contour correlation
- third concept using aerea momentum.

For all three concepts that part of the image containing the target is evaluated. Methods of image restoration may be applied if the image is corrupted by noise. If the results of evaluating the image are represented to the observer in an interactive mode methods of image enhancement may be used. The next step in processing the image is to digitally extract edges and special characteristics out of the image which are used for interpreting geometrical structures and to find the contours of the target image. These features can be detected by the contrast to the background or to other structures. For IR-images contrast in the image are caused by temperature differences. Therefore leading edges of the wings and the tail, the fuselage nose and the engines can be detected.

The first concept of pattern recognition uses aspect-angle-invariant length ratios. Aspect-angle invariance means that no absolute lengths or angles should be evaluated. The basic idea is to only evaluate length ratios which reproduce themselves for any projection angle as shown in figure 9. Using the center fuselage line (which is defined in the image by the line from the fuselage nose to the center of the engine outlets) as reference line one can define such length ratios for example:

- ratio of length of the line from the intersection point of the leading edges of the wings (or tail wings) to the engine outlets compared to the length of the reference line
- ratio of the length of the line from the air inlets to the engine outlets compared to the length of the reference line.

Taking the wing span as a second reference line one can define additional length ratios for example:

- ratio of tail wing span compared to the wing span.

These length ratios are evaluated in advance for the aircraft types to be identified and stored in a catalog. In the image of the target the detected edges are interpreted in their meaning by using geometrical information of aircraft structure, for example that the intersection of the leading edges of the wings are near by the middle of the fuselage or that the ends of the leading edges of the wing have by symmetrical reasons (also in any projection) the same distance from the fuselage reference line. One has to recognize that this is not an easy interpretation regarding for example "normal" aircrafts of the type delta-canard.

After identifying the meaning of the edges the length ratios as mentioned before can easily be deducted. By comparing these length ratios with the stored data in the catalogue the special type of the aircraft can be identified with some significance.

In figure 10 an aircraft of the type F-18 is shown. From this picture the length ratios a_i ($a_i = A'_i/L'$. i = 1,3,6,7 and $a_9 = A'_9/S'$, ' means projection) are deducted. The same parameters were derived in advance from aircraft views of MIG-25, F-15 and F-18. Evaluating the criteria, as shown in figure 11, (a_j^k means parameter of the catalogue aircraft k) the aircraft of figure 10 is identified with significance as a F-18 /2/.

The second concept of pattern recognition for aircraft identification is the correlation of the contour of the target aircraft with stored contours of catalogue aircrafts. The image must be processed for finding the outer contour of the target aircraft. In a higher contrast environment this may not be a difficult task. But one has to take into account the effects of shadows in a TV-image and temperature contrast within the contour in an IR-image. The contour is then extracted in the gray-shades, white and black. This contour has to be correlated to the contours of catalogue aircrafts. It seems to be convenient to store the catalogue aircrafts as three-dimensional structures. For each catalogue aircraft a number of cross-sections of the fuselage are stored. The wings are described by the corners of polygons where symmetrical structures should be stored only for one part. For the correlation of the contours the projection of this stored three-dimensional structure must be calculated by means similar to "computer aided design". This calculation has to take into account the distance of the target aircraft to get the right size of the projected contour and the aspect angle to get the right projection. If the aspect angle is not known very exactly one may calculate the contours for slightly modified aspect angles. These calculated contours of the catalog aircraft then must be correlated with the contour of the target aircraft. This can be performed as an area correlation for two images with one bit gray shade representation. The maximum of the correlation values of the contours of all catalog aircraft identifies with some significance the target aircraft.

The third concept is discussed in more detail in the following chapter.
Besides the above mentioned approaches in the literature a variety of methods are described, with more or less reference to their general practical usefulness for our applications. For computer implementation and tests a bearing concept had to be established.

5. PROPOSED METHOD SUPPORTED BY COMPUTER SIMULATION RESULTS

5.1 Assumptions and requirements

For practical studies supported by computer implementation and tests a method had to be selected which would supply the following properties.

- few computational work and storage (real-time processing!)
- sensitive for structural perturbations (global geometric properties) (high degree of separability of aircraft types)
- rugged against local noise and local geometric perturbations (weapon, optical effects, filtering errors, edge errors)
- translatorical and scale invariance (sufficient to a certain degree)

These properties promised to be provided by the well-known momentum method used in the OCR. Thus the algorithms described in this chapter and initially implemented on a minicomputer are based on this method with respect to the central identification process. The main results of some tests are presented in this chapter and further works in this direction are proceeding.
The algorithms were developed under the assumption of the availability of range and target orientation data. The identification process here described works on the premises that the assumed relative target position is roughly known (a X-Y-window in the image plane with a possible complete target anywhere in it) and that the relative aspect angles are known with sufficient accuracy. This information could for example be extracted by a radar-based target tracker.

The following aircraft types were used as target data base:
F-14 A, F-16 A, SU-15, F-18, Tornado, SU-11, MIG-23, MIG-25 (cf. figure 12, 13).
The target size was varied near 10 pixels per 0.5* target dimension and background noise was generated by simulating a Gaussian random noise with uncorrelated error probability of 0.1 to 0.2 per background pixel.

5.2 Description of used pattern recognition methods

PREMISES

The flowchart shown in figure 14 assumes the digitized image of a complete target within a "target window" sized b_z x h_z pixel and stored in the core of the digital identification computer. Furthermore it is assumed that time-correlated distance and target orientation are stored.

IMAGE PREPROCESSING

Noise abatement within the target window is done by an appropriate two dimensional filter (window threshold function) and as a result a preprocessed image is stored for further processing. Since small sized target patterns or target pattern elements may suffer local contour pertubations the proper identification process must be sufficiently insensitive to, without loss of sensitivity to the essential distinctive geometric feature.

IDENTIFICATION PROCESS

The identification process provides a mapping

$$I : G (b_z; h_z; x,y) \longrightarrow R^n$$

from the set of gray value functions over the target window b_z, h_z into an n-dimensional signature space R^n. n should be chosen such that a sufficient separation of the target catalogue is guaranteed using only reliable and efficient criteria for separation. In our works we have suggested to make use of the momentum method (cf. /7/, /8/. /9/. /10/). Setting

$$B (x,y): = \text{gray value function of a pixel } (x,y)$$
$$\text{within the target window } Z = b_z \times h_z$$

the central moments $\mu_{i,j}$, i, j e N_0 are calculated:

$$\mu_{i,j} = \iint_Z (x-\bar{x})^i \ (y-\bar{y})^j \ B(x,y) \ d(x-\bar{x})d(y-\bar{y})$$

$$\bar{x} = \iint_Z xB(x,y)dxdy \ / \ \iint_Z B(x,y)dxdy \qquad \bar{y} = \iint_Z yB(x,y)dxdy \ / \ \iint_Z B(x,y)dxdy$$

For separation, the specific target area (F (b_z, h_z) = μ_{oo}/r with respect to the precalculated target distance r and two algebraic invariants $G_1(b_z, h_z)$, $G_2(b_z, h_z)$ based on $\mu_{i,j}$, i + j 2 are computed. These threee criteria are translatorically invariant and rotational invariant with respect to the imaging plance. G_1 and G_2 are (theoretically) scale invariant without using r. In practice discretisation errors may cause some trouble especially for small sized target patterns.

Due to translatorical invariance it is not necessary to find out some "geometric center of gravity" or any other reference point or line for exact location of the target relative to the origin of the X-Y-image plane.

The rotational invariance may be used for storage reduction of the target catalogue. The invariants F, G_1, G_2 of the target window $b_z \times h_z$ are compared with the respective values of the aircrafts stored in the target catalogue selected via the precalculated target orientation. Special aircraft properties like variable wing geometry affect the target catalogue in that way that for a fixed orientation and aircraft type the characteristics of the aircraft are mapped on a tracetory in the signature space R^3 whereas fixed wing aircrafts are mapped on a single point.
By means of a simple distance criterion in that space the most probable aircraft type is computed. This criterion provides an additional information about possible alternatives and the reliability of the identification result.

5.3 Identification examples

The described algorithms were written in FORTRAN and together with the target catalogue (cf. figure 12 and 13) implemented on a minicomputer.

The first identification example (F-18 side view) is to be seen in figure 15a and 15b. The identification result is shown in figure 16. The area and shape invariants F, G_1 and G_2 provide the correct identification result even without noise filtering. G_1 and G_2 offer the F-18 and SU-11 as possible candidates. The area criterion however excludes the SU-11 so that the F-18 remains as correct result.

Figure 15c and 15d shows as second example the top view of a F-18 before noisy background and the digitally filtered target window.

Figure 17 again shows the identification result. The shape invariants G_1 and G_2 leave as possible alternatives the F-16 A, F-18 and MIG-25. The area criteria F again correctly identifies the F-18.

5.4 Outlook

The proposed basis algorithms for target identification are being subjected to extensive efficieny tests under different environmental considerations. Special effort is applied to general algorithms for separation of target areas and background areas under practical conditions (i.e. seriously perturbed incoherent

IR-image elements, effect of aspect angle errors due to measurement by other sensors like tracking radar).

6. CONCLUSIONS

In aircombat situations the operator task performance can be improved by multi-sensor-based and computer-aided target identification. It has to be considered that in aircraft application the computer capacity is limited and that the identification process has to be done in nearly real time. The proposed methods meet the requirements for aircraft implementation, and the first results of our investigations are very encouraging.

The investigations are to be extended, to determine the dependence of the aspect angle and of the impaired aircraft contour due to weapon payload.

REFERENCE

/1/	:	Aimval / Aceval shows need for electro-optics. Perspective, Feb. 1978, pp 61-63
/2/ ESG	:	Sensorverbundsystem für Kampfflugzeuge, ESG-Report 5400-BT-0001 (unplublished).
/3/ W. Kuhl	:	Fernsehaufnahmeröhren für niedrige Strahlungsleistungen in LLL TV-Systemen Valvo Berichte, Band XVII, Heft 1, Sept 1972
/4/ ESG	:	Elektro-Optischer Pod für den TKF-Einsatz im Luftangriff ESG-Dok-Nr.: TKF-BT-0005, Teil 1 vom 31.1.1979
/6/ RCA	:	Electro-Optics Handbook
/6/ H.P. Baier et al	:	Die Erkennbarkeit von Landoldtringen im Fernsehbild bei Bandbegrenzung und gleichzeitigem rauschförmigen Störsignal
/7/ Ming-Kuei Hu	:	Visual Pattern Recognition by Moment Invariants, IRE Trans.Inform. Theory, Vol. IT-8, pp. 179 - 187, 02/1962
/8/ Niemann H.	:	Methoden der Mustererkennung, Akademische Verlagsgesellschaft, Frankfurt 1974
/9/ Kugler J. Wahl F.	:	Kantendetektion mit lokalen Operatoren, Angewandte Szenenanalyse, DAGM Symposium Karlsruhe 1979
/10/ Winkler G. Vattrodt K.	:	Maße für die Auffälligkeit in Bildern, Bildverarbeitung und Mustererkennung DAGM Symposium Oberpfaffenhofen 1978
/11/ ESG	:	Sensorverbundsystem für die taktische Identifikation im TKF ESG-Report: TKF-BT-0503
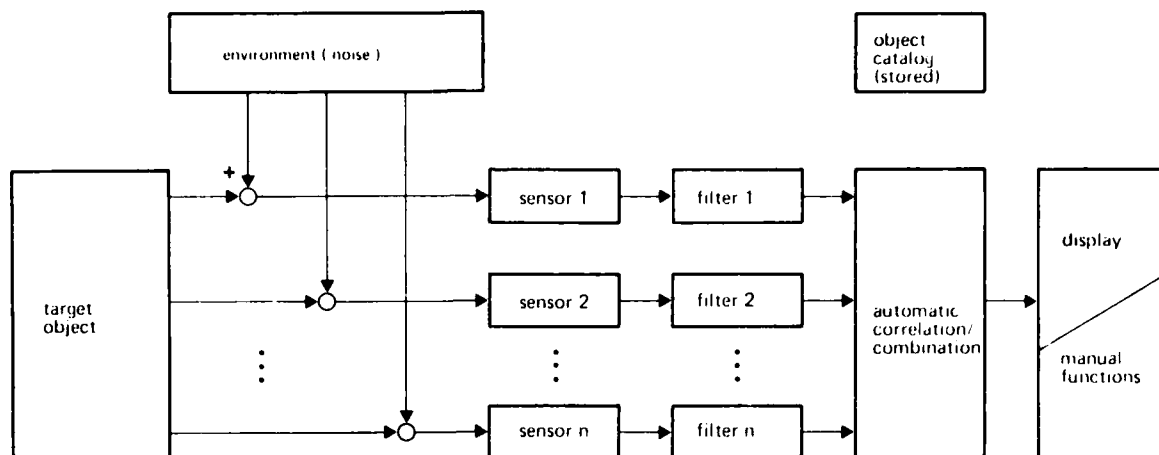
Fig. 1   Multi-sensor-approach for target identification (sensor fusion)
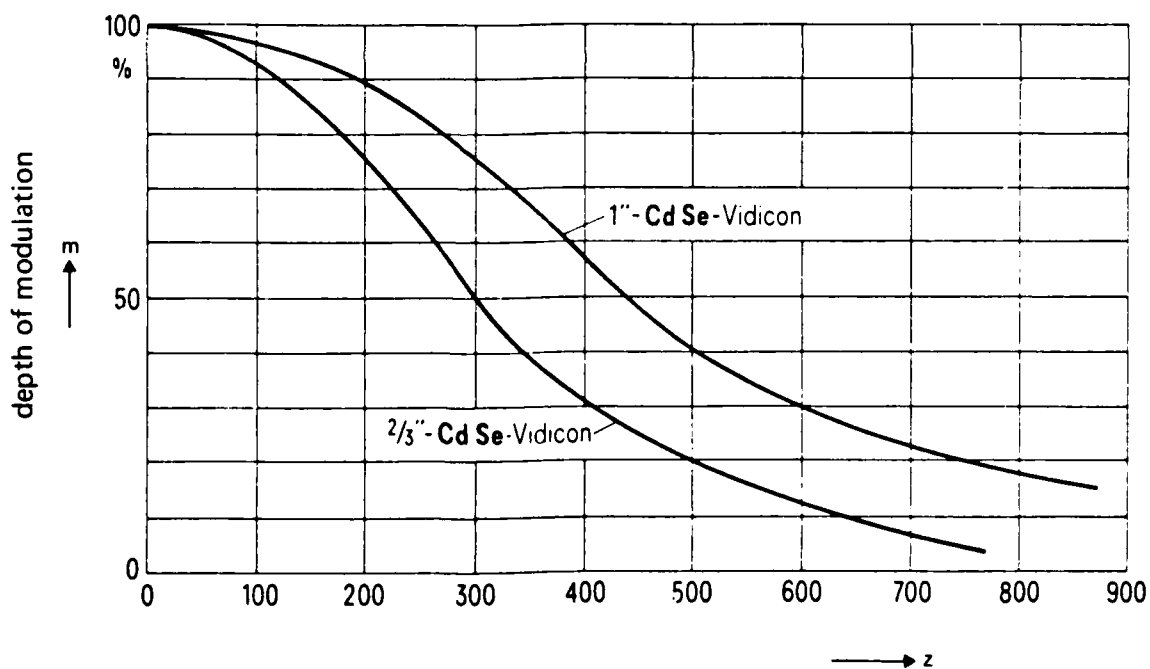
Fig.2   Depth of modulation as a function of the number of lines z

$$E = 210\,\mu A/lm \qquad U_K = 10\,kV \qquad U_T = 10\,V$$
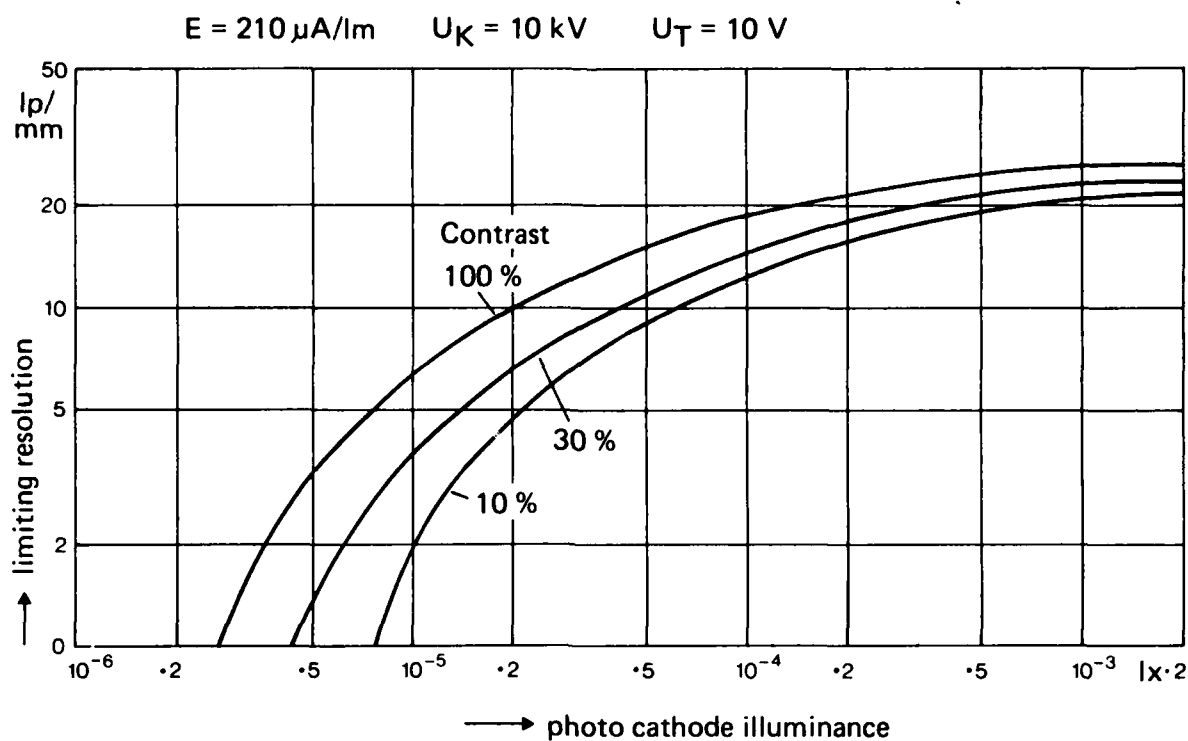
Fig. 3 Limiting resolution as a function of photo cathode illuminance

Diameter of Optic D [mm]	Wave-Length [μm]			
	0,5	1	4	10
50	0,012	0,024	0,1	0,24
100	0,006	0,012	0,05	0,12
150		0,008	0,03	0,08
200			0,024	0,06
250				0,05

$$\delta\varphi = 1.22 \cdot \lambda / D$$

Fig. 4 Theoretical angular resolution for diffraction limited optics (in mrad)

33-10
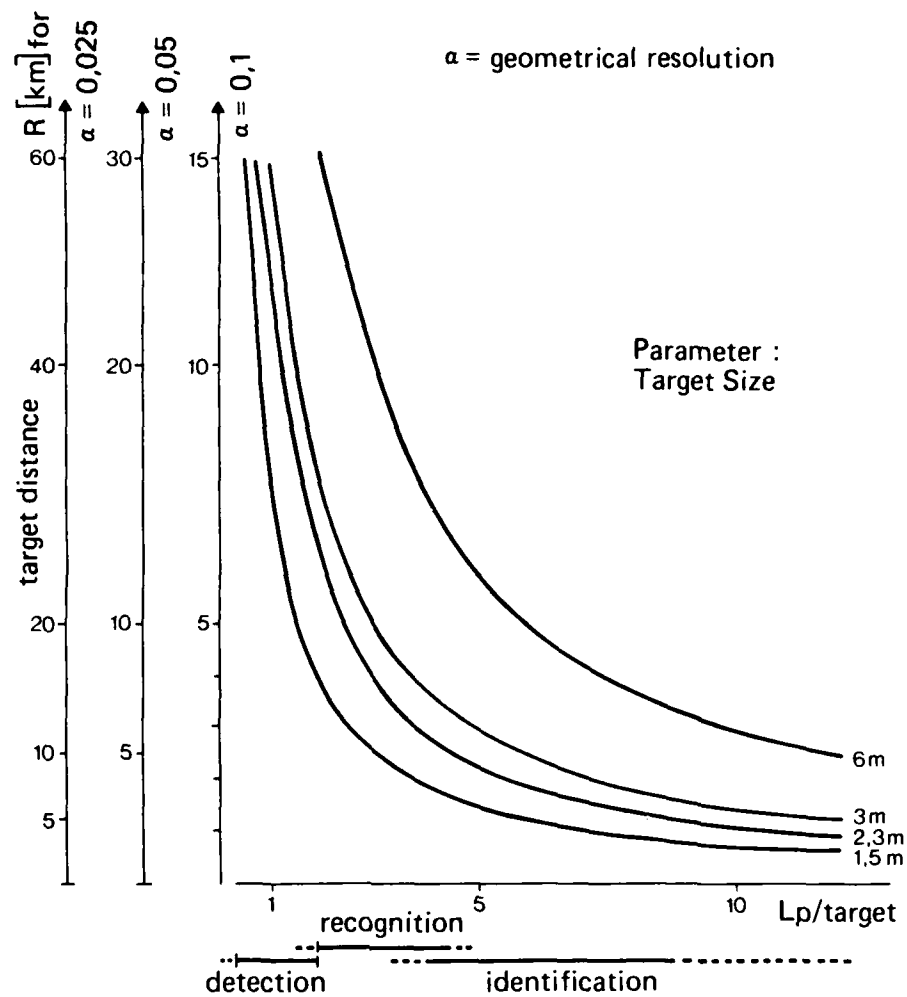
Fig. 5   Sensor range due to the geometrical sensor resolution
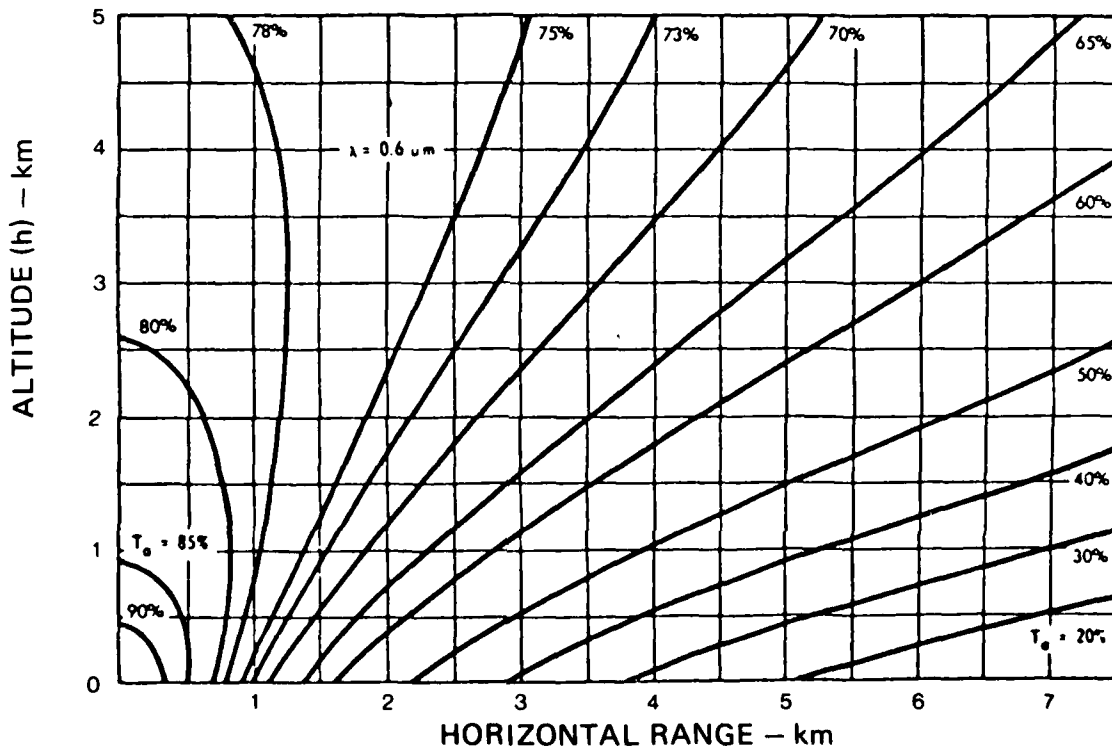
Fig. 6   Contours of constant atmospheric transmittance for radiation at 0,6 µm in a atmosphere-visibility 12 km
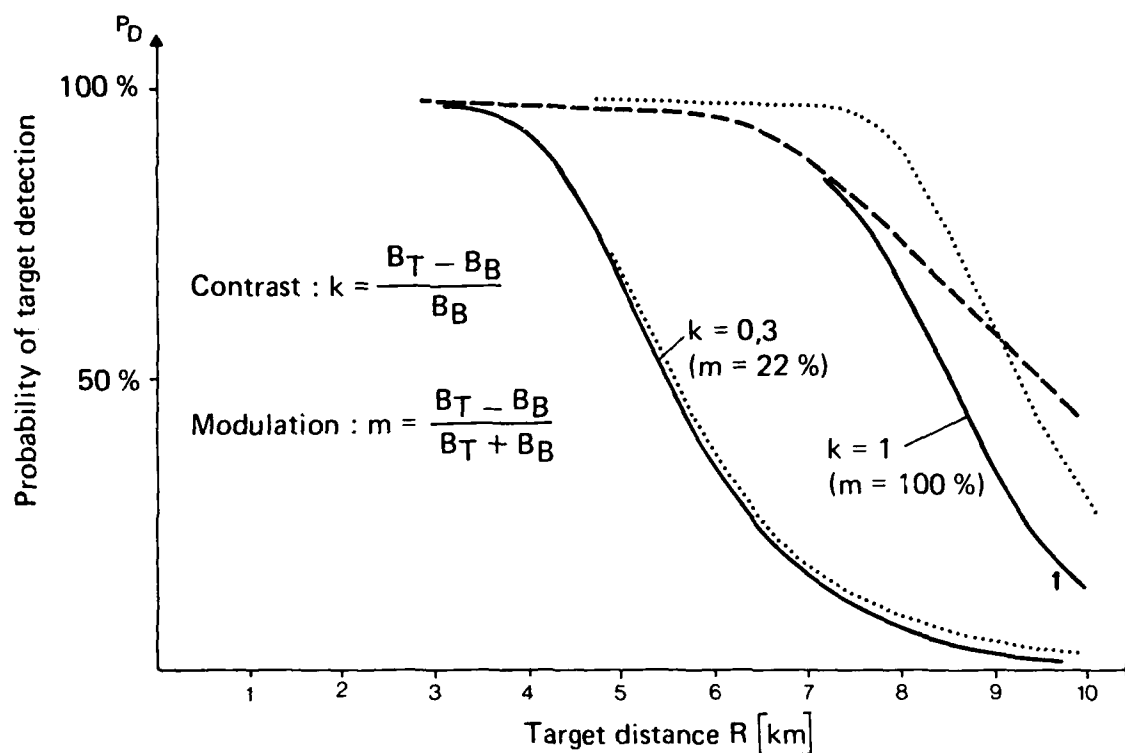
Fig. 7   Target detection with TV-System (FOV : 2° x $2,65^o$)

Contrast : $k = \dfrac{B_T - B_B}{B_B}$

Modulation : $m = \dfrac{B_T - B_B}{B_T + B_B}$

$k = 0,3$
$(m = 22\,\%)$

$k = 1$
$(m = 100\,\%)$

Parameter : Field of view : 2° x $2,65^o$

Target : $\Delta T = 2K$
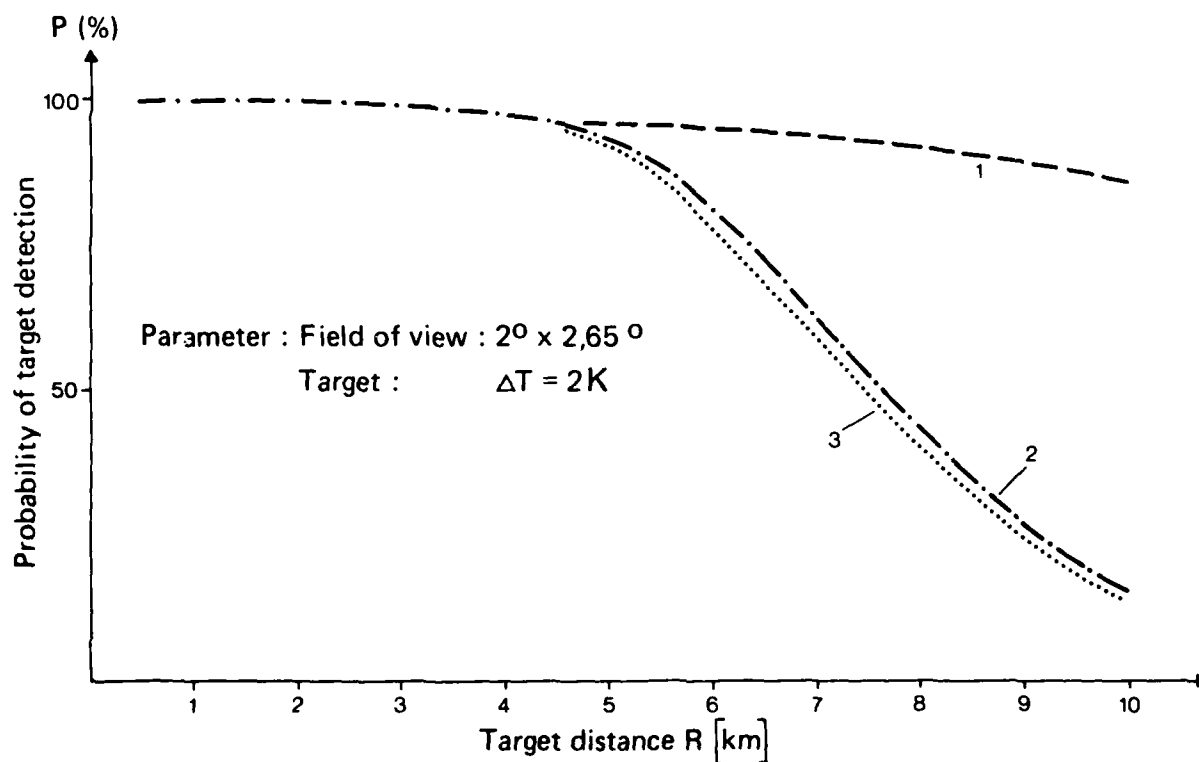
Fig. 8   Target detection with FLIR (FOV : 2° x $2,65^o$)

projection

$$\frac{a}{b} = \frac{c}{d}$$
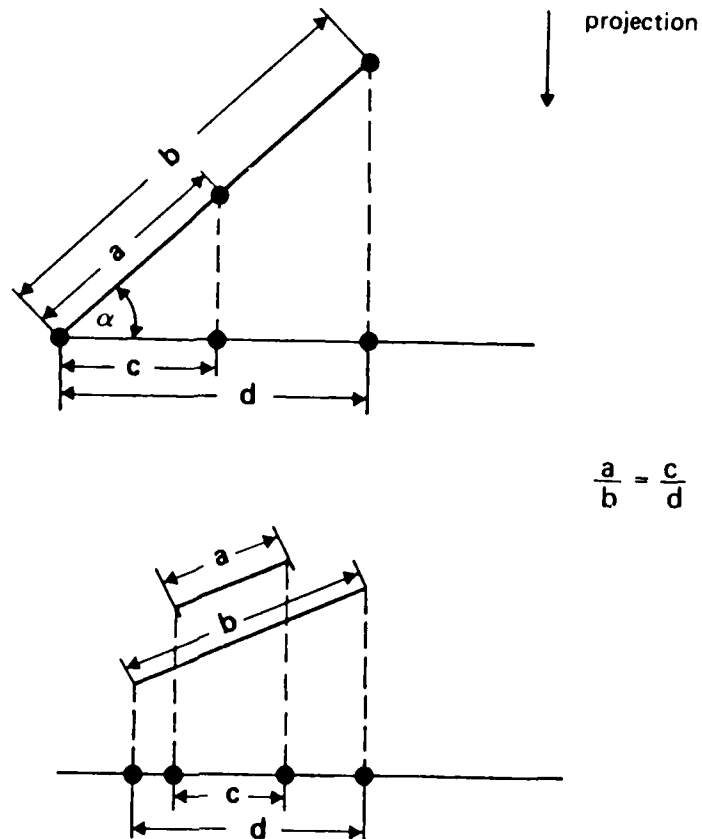
Fig. 9   Reproduction of length ratios by a projection
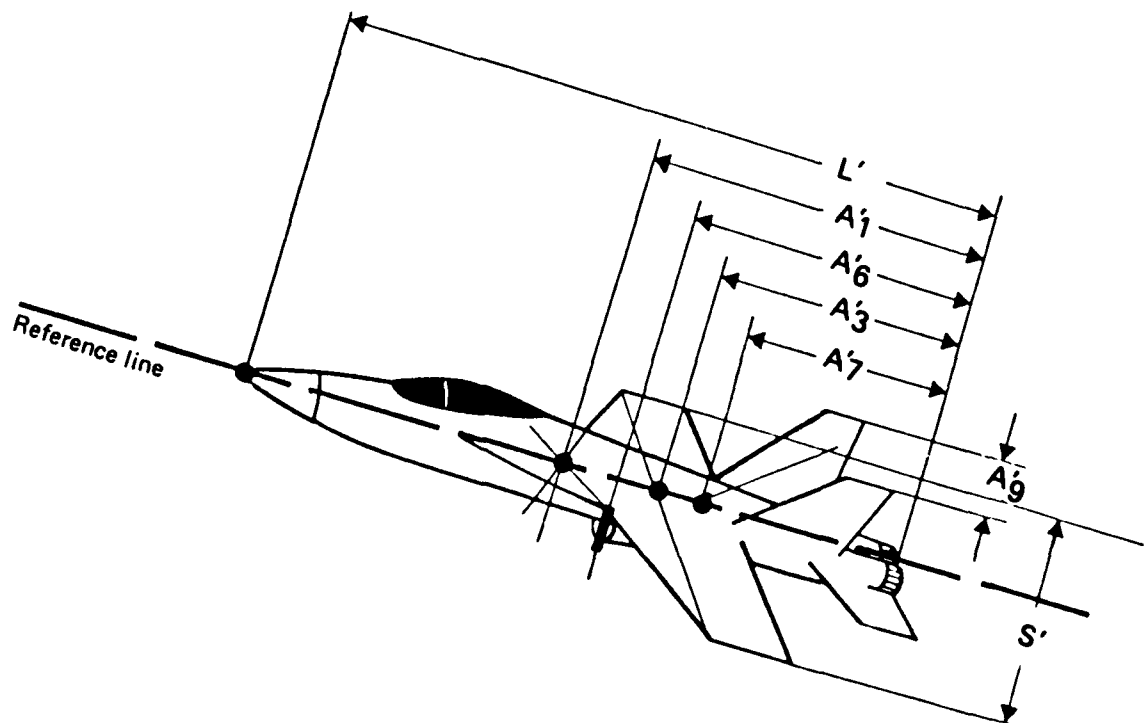
Fig. 10   View of a F - 18 with length measures

criterion J_v^k	MIG 25	F - 15	F - 18
$/ a_1^k - a_1 /$	0,02	0,06	0,03
$/ a_3^k - a_3 /$	0,09	0,15	0
$/ a_5^k - a_5 /$	0,19	0,19	0,09
$/ a_7^k - a_7 /$	0	0,13	0,01
$/ a_9^k - a_9 /$	0,02	0,03	0,02
$J_v^k = - \Sigma$	− 0,32	− 0,56	− 0,15

Fig. 11 Evaluating the criterion for identifying the aircraft of Fig. 3
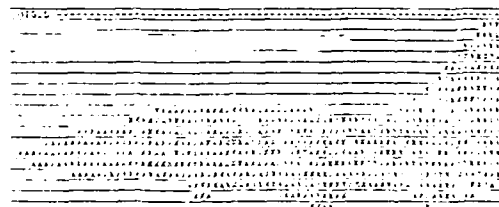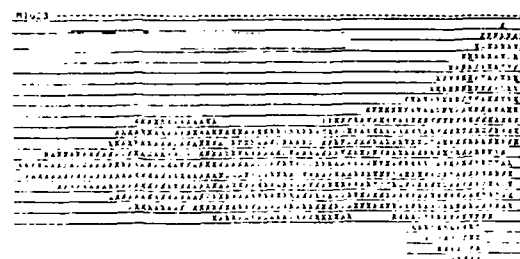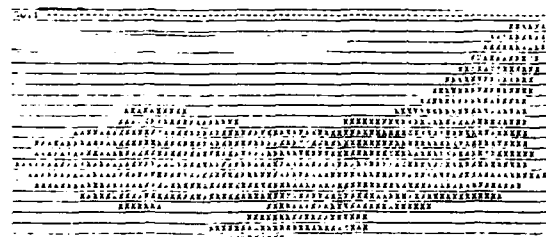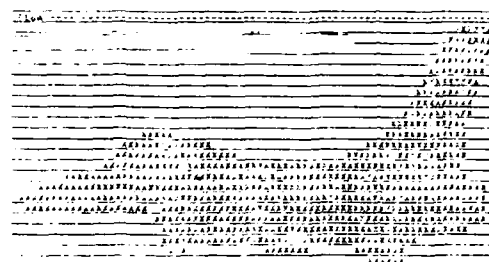
Fig. 12 Target catalogue (side views, aircraft sizes not in scale)
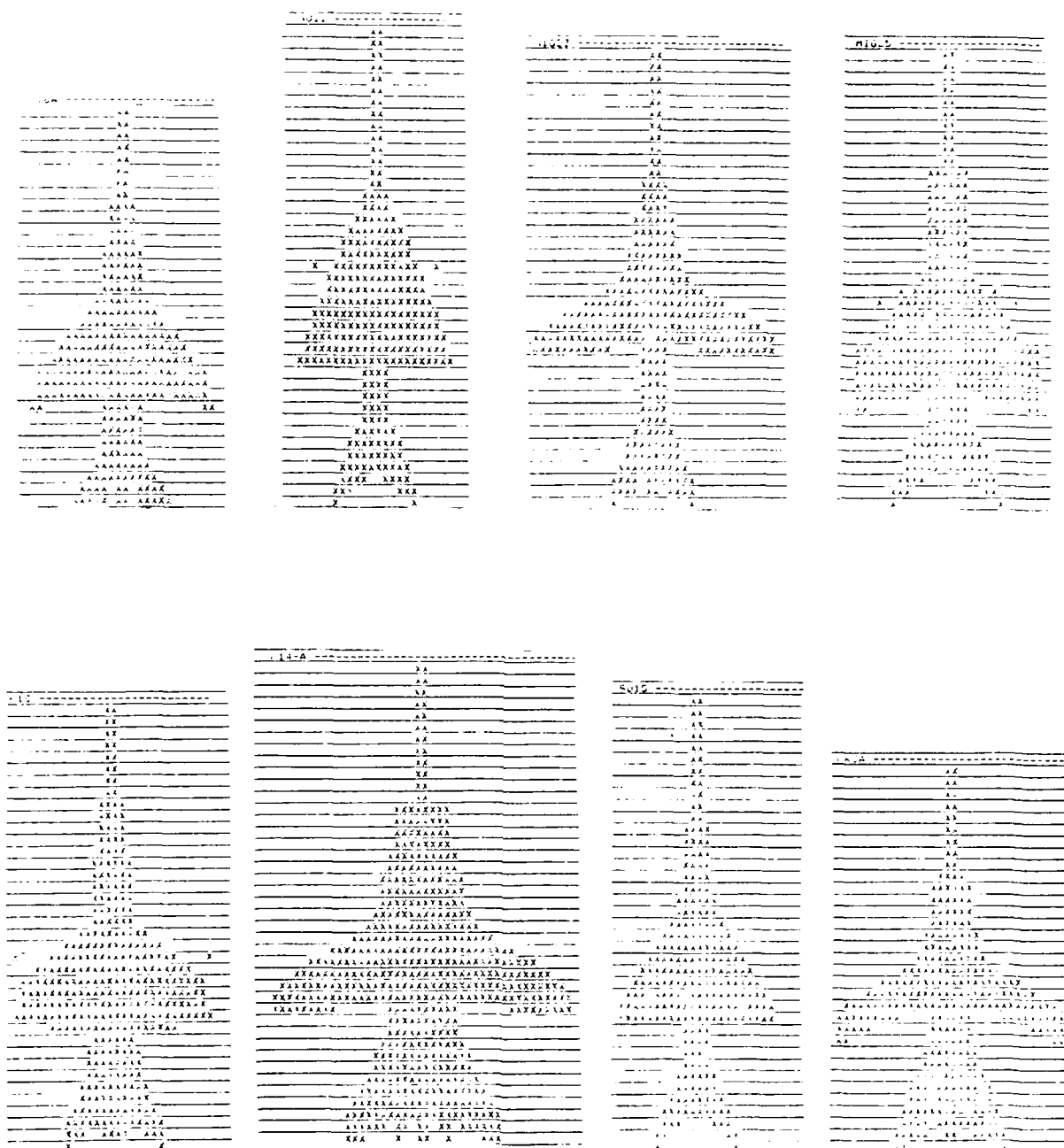
Fig. 13 Target catalogue (top views, aircraft sizes not in scale)

33-16
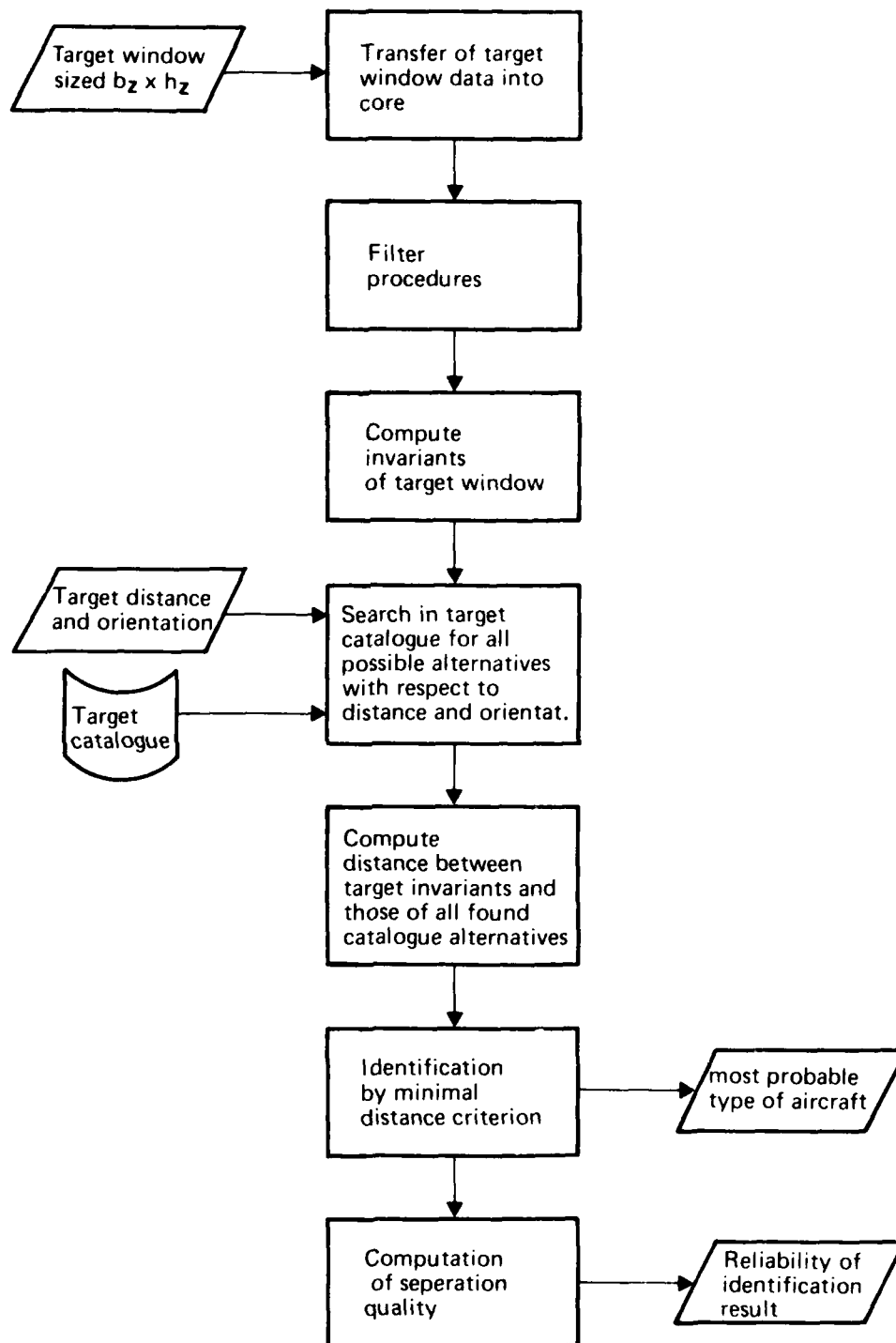
Fig. 14 Identification algorithm

Fig. 15 Noise Filtering

a) Algebraic invariants for identification

b) Area criterion for identification
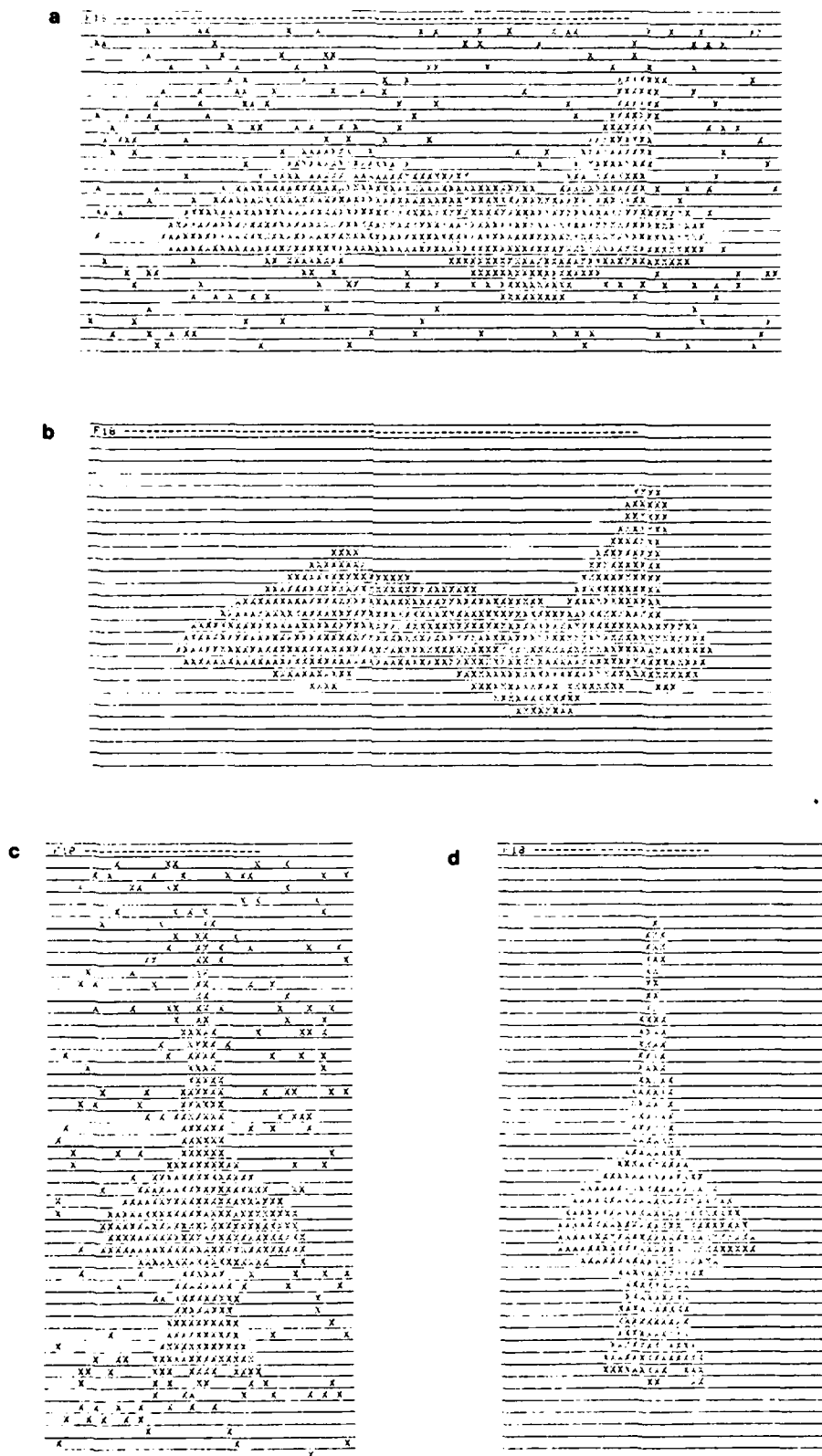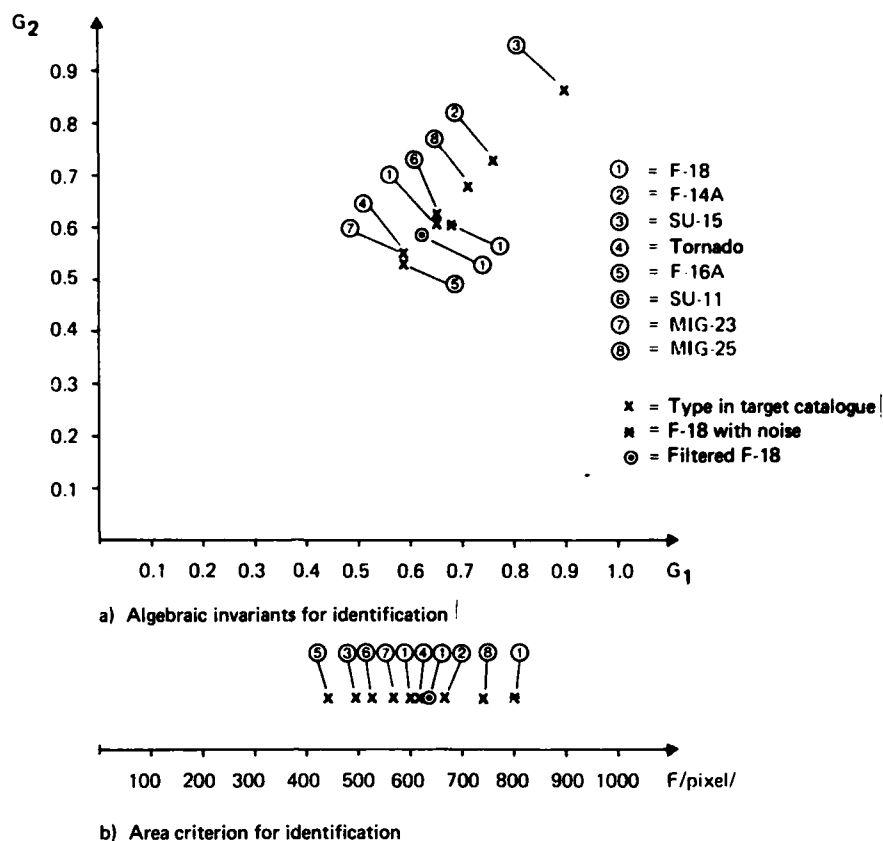
Figure 16 : Identification of a F-18 (side view) according to Figure 15a and b and target catalogue

a) Algebraic invariants for identification

b) Area criterion for identification
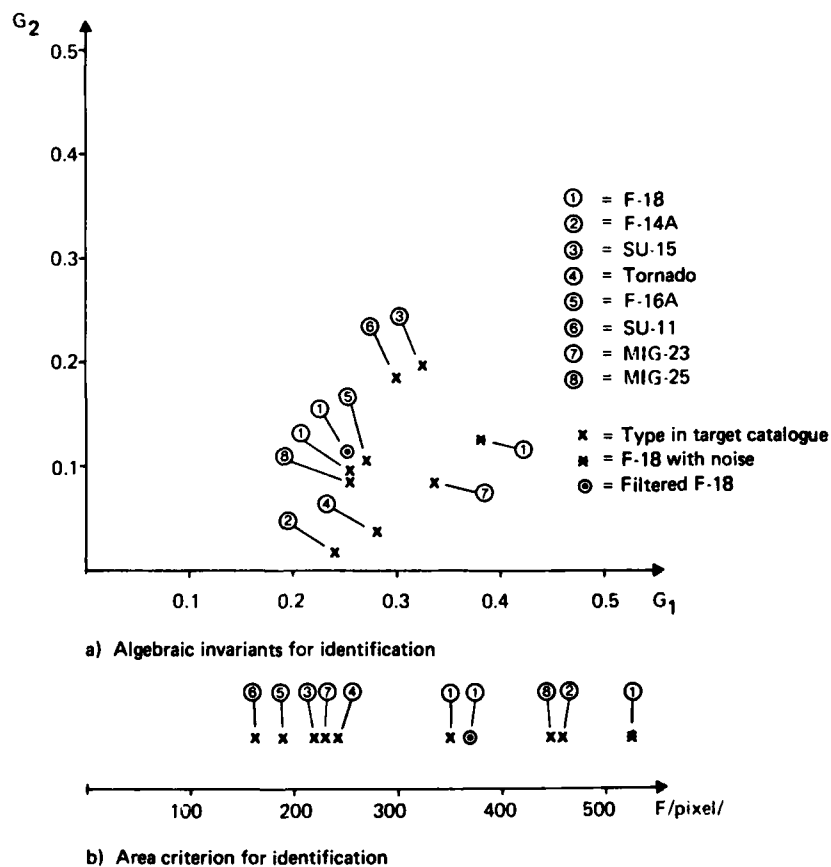
Figure 17 : Identification of a F-18 (top view) according to Figure 15c and d and target catalogue

AUTONOMOUS TARGET SCREENERS
BY
R.K. AGGARWAL, M. GEOKEZAS, D.E. SOLAND
SYSTEMS AND RESEARCH CENTER
HONEYWELL, INC.
2600 RIDGWAY PARKWAY
MINNEAPOLIS, MINNESOTA 55413

SUMMARY

Over the past ten years Honeywell has been involved in the development of infrared imagery target screeners under internal, AFAL, NV&EOL and DARPA support.

The sophistication of reconnaissance and strike systems is continually increasing due to the high threat operational environment. Thus, advanced Forward Looking Infrared (FLIR) sensors are integrated on high performance aircraft. The task loading and high information rate of advanced sensors has made it impossible for a human to perform the target search/detection/recognition task accurately, consistently, and in real time.

The judicious application of image enhancement, automatic control functions and target screening will improve the sensor/operator interface and lead us closer to the realization of an autonomous target screener.

In this paper we describe an autonomous target screener concept. The basic functions of an autonomous target screener are, segmentation, feature generation, classification (detection/recognition), and symbol generation.

Image segmentation is the function by which the image is segmented in background and objects of interest. The image information within these objects of interest is processed to generate a set of features which characterize the targets of interest. The classification function utilizes a statistical/syntactic classifier for detection (target vs. clutter decision) and recognition (truck, tank, APC, etc.). A symbol indicating the position and type of target is displayed on the monitor for cueing purposes.

1. INTRODUCTION

Since 1969, Honeywell has been involved in research and development of infrared imagery target screeners with internal, Air Force Avionics Laboratory (AFAL), Night Vision and Electro Optics Laboratory (NV&EOL), and Defense Advanced Research Projects Agency (DARPA) sponsorship.

The high threat operational environment necessitates increased sophistication of reconnaissance and tactical systems. Thus, advanced Forward Looking Infrared (FLIR) sensors are integrated on high performance aircraft. The high information rate of advanced sensors and task loading has made it impossible for human operator to perform the target search/detection/recognition task accurately, consistently, and in real time.

The judicious application of image enhancement, automatic control functions, and target screening is expected to improve the sensor/human interface and eventually resulting in the realization of an autonomous target screener.

Different approaches, generic concepts, laboratory tests, and conclusions on autonomous target screeners are presented in this paper.

2. AUTONOMOUS TARGET SCREENER

The basic functional block diagram of the future Autonomous Target Screener System (ATSS) is presented in Figure 1. It consists of a Forward Looking Infrared (FLIR) imaging sensor, an image enhancement module, a target screener and a TV compatible display.

The FLIR with its standard 525/875 lines per frame, 30 frames per second format and TV compatible display will not be discussed here. Image enhancement and target screener will be discussed next.

2.1. Image Enhancement

Image enhancement for visual TV and FLIR has been and continues to be fertile grounds for research and development. FLIR image enhancement may consist of D.C. restoration, responsitivity equalization, resolution and super resolution, minimum resolvable temperature (MRT), and contrast enhancement. The first four improve the image quality and have been discussed elsewhere (Narendra, P.M., 1977). The last optimizes the interface between the FLIR sensor and display and will be discussed in this paper, because it is an integral part of an autonomous target screener and has been reduced to practice (Narendra, P.M., 1978).

2.1.1 Contrast Enhancement

The scene dynamic ranges (1000:1) encountered by imaging sensors can be much higher than the CRT display dynamic range of (20:1). In addition, the scene intensity extrema are changing continuously.

In order to match the imaging sensor output to the display input we need a global automatic gain/bias control (AGC) as well as a local area contrast enhancement. The first determines the compression/expansion for scene images to fit into the display dynamic range and the second enhances the local contrast.

The adaptive contrast enhancement algorithm performs the following functions:

- Vary the local average brightness (bias) so that overall dynamic range of scene is compressed;

- Enhance local variations above the contrast sensitivity threshold of the human eye; and

- Automatically fit the intensity extrema in the enhanced image to the display limits

A functional block diagram of this algorithm is shown in Figure 2. The image intensity at each point is transformed based on local area statistics -- the local mean m_{ij} and the local standard deviation σ_{ij} are computed on a local area surrounding the point.

The local area mean is first subtracted from the image at every point. A variable gain is applied to the difference to amplify the local variations. A portion of the local mean m_{ij} is then added back to restore the subjective quality of image. The local gain G_{ij} is adaptive, being proportional to M, to satisfy psychovisual considerations (Weber's law); and inversely proportional to σ_{ij}, so that areas with small local variance receive larger gain.

The transformed intensity is then:

$$I_{ij} = G_{ij} (I_{ij} - m_{ij}) + m_{ij}$$

where, the local gain $\quad G_{ij} = \alpha \dfrac{M}{\sigma_{ij}} , \quad 0 < \alpha < 1$

where M is the global mean.

To prevent the gain from being inordinately large in areas with large mean and small standard deviation, the local gain is actually controlled as shown in Figure 3.

2.2   Target Screener

The target screener processes the raw or enhanced FLIR imagery and displays a symbol on the display indicating the type and location of detected targets. The screener serves as a cuer to the operator who then makes the final decision.

The need for an automated target screener arises mainly from the task loading and high information rate of advanced sensors. The human operator is hindered in performing his task by display and human limitations. We discussed in Section 2.1.1 how contrast enhancement results in a hands-off display operation. Humans are limited by physiological and psycological factors in a threat rich environment, resulting in low probability of detection and recognition. Under ideal laboratory conditions, human performance has been reported (Krebs, M.J., 1974) in terms of probability of detection; P_D = 70% and recognition; P_R = 50%, with corresponding search average times of 4.5 and 5.5 seconds. Under field tests (STANO, 1971). the acquisition time increases and the probabilities of detection and recognition decrease.

An autonomous target screener consists of the following four functions; image segmentation, feature extraction, classification and symbol generation as shown in Figure 4. Each of these functions will be discussed below.

2.2.1   Segmentation

Segmentation is the function by which the image is segmented in background and objects of interest. There are basically two approaches to segmentation; one based on edge and intensity thresholding and the second is region based. These approaches are discussed below.

2.2.1.1 Edge and Intensity Thresholding

This approach processes the FLIR image through an edge filter, determines the intensity and edge statistics, estimates local or global thresholds and segments the image. This approach is called autothreshold and adapts the edge and intensity thresholds to changing scene contrast and intensity levels (Panda, 1978).

Figure 5 shows the overall concept for autothreshold. Briefly the function of each box is as follows: Raw video is passed through a low-pass smoothing filter. This limits the bandwidth of the noise. The smooth data is an input to both the edge filter and the bright filter. The edge filter generates the edge magnitude of the smoothed data from which an edge threshold is determined for the next scan line. The output "Edge" is a binary signal obtained by comparing the analog edge signal with the edge threshold. The bright filter determines the background. The background image is processed to estimate the bright threshold and subtracted from the raw image. Comparing the latter with the intensity threshold the binary "bright" signal is generated.

The smoothing filter is a weighted average filter based upon the following equation

$$I(i,j) = \quad [A(i-1,j-1) + 2A(i-1,j) + A(i-1,j+1)$$

$$+2A(i,j-1) + 4A(i,j) + 2A(i,j+1) \quad (1)$$

$$+A(i+1,j-1) + 2A(i+1,j) + A(i+1,j+1)]$$

where $I(i,j)$ is the smoothed output at location (i,j) and $A(i,j)$ is the video intensity. This filtered video is then the input to the edge filter and to the bright filter. Figure 6 is a sample of five scan lines of FLIR video over a tank. Figure 7 is the resultant smooth data.

The two-dimensional edge filter (Sobel) calculates:

$$H_{j-1} = I(i-1,j-1) + 2I(i,j-1) + I(i+1,j-1)$$

$$H_{j+1} = I(i-1,j+1) + 2I(i,j+1) + I(i+1,j+1) \quad (2)$$

$$V_{i-1} = I(i-1,j-1) + 2I(i-1,j) + I(i-1,j+1)$$

$$V_{i+1} = I(i+1,j-1) + 2I(i+1,j) + I(i+1,j+1)$$

where H and V are the horizontal and vertical components. Then the edge value associated with (i,j) pixel is

$$E(i,j) = |H_{j+1} - H_{j-1}| + |V_{i+1} - V_{i-1}| \quad (3)$$

Figure 8 is the Sobel edge for the data shown in Figure 7. Superimposed on the edge data is the adaptive edge threshold. The edge threshold is

$$E_n = K * \overline{E_{n-1}} \quad (4)$$

where E_n is the threshold, $\overline{E_{n-1}}$ is the previous scan line edge average and K is an optimum constant statistically determined.

As was mentioned previously, the smoothed video is fed into the bright filter. The primary function of the bright filter is to estimate the background intensity. This background estimate is continually updated as the image is scanned by a recursive filter and updating logic. We determine if there is large contrast between scan lines on a pixel by pixel basis. The background estimate J is built up over several scan lines. The background is updated from scan line to scan line and is defined as

$$J(i,j) = J(i-1,j) + (1-\beta) I(i,j) \quad (5)$$

where β is a constant and $0<\beta<1$. When not updating $J(n,k) = J(n-1,k)$. Once the background $J(i,j)$ is determined, it is subtracted from $I(i,j)$ to give a zero reference. Hence, $Z(i,j) = I(i,j)-J(i,j)$ is data with zero reference to be thresholded. $Z(i,j)$ is compared to the bright threshold EPSI defined as

$$EPSI = \frac{\mu}{Num(1-\beta)} \quad |J(i,j) - J(i-1,j)| \quad (6)$$

where μ is a constant. The output of the comparison is the Bright signal and includes hot and cold objects. Figure 9 shows the background estimate for the smoothed video (Figure 7). The resultant video to be thresholded is shown in Figure 10 with the threshold superimposed. The thresholded edge and intensity signals are combined in a logical way to extract an interval over which an object may exist. The presence of a leading edge, followed by a bright and concluding with a trailing edge constitutes an interval. If the interval width falls between predefined max/min limits, the scan line, starting column, and interval width are stored.

Accumulating congruent intervals, candidate objects are extracted. Figure 11a represents raw FLIR image. Figure 11b represents the binary edge, Figures 11c and 11d the hot and cold binary intensities, Figure 11e the object intervals and Figure 11f the extracted objects.

2.2.1.2 Prototype Similarity

The image segmentation scheme using prototype similarity transformation (Aggarwal, R.K., 1978) can be divided into the following major steps:

- Attributes
- Prototype Generation
- Threshold Selection
- Prototype Inference
- Cell Inference
- Similarity Relation

STEP 1: Attributes. A cell represents a single pixel or a collection of pixels depending upon the required resolution in the segmented scene. Some of the commonly used attributes are average intensity, edge texture, etc. Suppose $x_1,...,x_N$ are the N

attributes characterizing each cell. These N attributes may be N independent measurements on each cell or may be N functions of M (M≥N) independent measurements.

STEP 2: Prototype Generation. For each of these N attributes characterizing a cell, a two-dimensional distribution function $F(j,i)$ is calculated as follows: Suppose the attribute value of a cell is i. Count the number of cells in some experimentally chosen neighborhood (depending upon the resolution, size of the target, etc.,) that have attribute value j. Accumulate this sum for all the cells in the picture that have attribute value i. This sum gives $F(j,i)$. Do this for all values of i and j.

Next initial background and target prototypes are determined using a priori information about the scene. This can be done by locating typical background and target cells or by using some attribute information about the background/target. For example, a running motor is the brightest part of tactical FLIR images.

Let the target cell attribute value be A_T and background cell attribute value be A_B. Based on these two values A_T and A_B, two intervals $|A_T*T_A, A_T/T_A|$ and $|A_B T_A, A_B/T_A|$ are calculated, where T_A is an empirically chosen threshold on the value of attribute A.

These two intervals are assumed to be disjoint. The case of overlapping intervals implies either a bad choice of target/background cues or a low value of threshold T_A.

These two disjoint intervals define the first two prototypes P_0 and P_1. For generating additional prototypes, consider the two-dimensional distribution function shown in Figure 12. All the cells that belong to prototypes P_0 or P_1 are zeroed (shown by hatched areas). Suppose the modified distribution function is $F'(j,i)$. Then for each attribute value i, we have an attribute profile of neighbors. By considering each value i in the intervals $|A_T T_A, A_T/T_A|$ and $|A_B T_A, A_B/T_A|$, the cumulative attribute profiles F_{P_0} and F_{P_1} are calculated as follows:

$$F_{P_0} = \sum_{i \in [A_T T_A, A_T/T_A]} F'(j,i) \tag{7}$$

$$F_{P_1} = \sum_{i \in [A_B T_A, A_B/T_A]} F'(j,i) \tag{8}$$

An example of these profiles is shown in Figure 13. A maximum is located in each of these profiles. The maximum of these maxima gives the location of the next prototype interval. This corresponds to maximizing the probability of finding a neighbor that has attribute value outside the attribute intervals of previous prototypes. Suppose the attribute value is A_2. This gives rise to an interval $[A_2 A_T, A_2/A_T]$ for the prototype P_2. At this stage, there are 3 prototypes P_0, P_1, and P_2. Now there are three cumulative profiles for the three intervals. The whole sequence of operations is repeated until no more prototypes can be generated.

STEP 3: Threshold Selection. A numerical value between 0 and 1 needs to be chosen for each attribute for defining prototype intervals. Too small a threshold leads to larger intervals and consequently fewer number of prototypes, whereas too large a threshold will lead to smaller intervals and larger number of prototypes. In the extreme cases, on one hand, we may have only two prototypes which will give rise to too many edge elements; and on the other, we may have many prototypes so that each cell is similar to only one prototype giving rise to too many different objects in the scene.

For FLIR images, a typical value for the number of prototypes for each attribute is somewhere between 10-15. So the thresholds can be adjusted to give the right number of prototypes.

STEP 4: Prototype Inference. Let $P_0,...,P_N$ be the set of prototypes. Each one of these prototypes has an interval on the attribute axis associated with it. Each cell in the picture is labeled by a string of prototypes it is similar to. A cell can be similar to more than one prototype as the prototype intervals can overlap. During the labeling process, a co-occurrence matrix is constructed. Each element A_{ij} in the co-occurrence matrix, i = 0,...,N; j = 0,..,N, corresponds to the frequency that the prototypes P_i and P_j occur together in labels.

The fact that the prototype P_0 was generated by a target cue and P_1 was generated by a background cue is used to infer meaning for other prototypes. The co-occurrence matrix is used to guide the inference. Suppose A_{0i} is maximum for $i = i_1$ and A_{1j} is maximum for $j = j_1$. Depending upon which one of A_{0i_1} and A_{1j_1} is greater, either prototype P_{i_1} or P_{j_1} is considered for inferring its meaning. The following rules are used to infer meaning for a prototype:

- A prototype whose interval overlaps a target interval and does not overlap a background interval is a target prototype.

- A prototype whose interval overlaps a background interval and does not overlap a target interval is a background prototype.

- A prototype whose interval overlaps both target and background intervals is an edge prototype.

- A prototype whose interval does not overlap target or background interval is assigned the "meaning unknown".

STEP 5: Cell Inference. Each prototype in a cell is replaced by its inferred meaning. The following string grammar is used to reduce string to a character:

```
TT → T
EE → E
BB → B
TB → E*
TE → T
BE → B
E*α→ E*  , αε {T,B,E,E*}
```

where

```
T => target cell
B => background cell
E*=> strong edge cell
E => weak edge cell
```

STEP 6. Similarity Relation. Based on each attribute, using the above described procedure, a meaning can be assigned to each cell of the picture. Thus each cell has a string of cell meanings, the length of the string being equal to the number of attributes needed. This is called a similarity relation. A cell should be assigned the same meaning by all the attributes before it is assigned that meaning. Otherwise, the cell is classified as "meaning unknown". A more complex relationship can be devised depending upon the type of imagery, type of attributes, etc.

The prototype similarity transformation was tried on FLIR images of tactical targets. The technique was first tried on full frames (520 x 480 pels) for target/background segmentation and then on the isolated targets for component extraction. The target center and its approximate size were recorded during digitization. The 8-bit digitized data was scaled down to 100 grey levels to cut the computer memory requirements for storing joint distribution function.

A cell was defined as 2 x 2 pels for component extraction and as 4 x 4 pels for target/background segmentation. A neighborhood of 3 x 3 cells was used for calculating the joint distribution function. A threshold of 0.85 was used for defining the attributed internals.

The results using the average intensity over the cell as the only attribute are shown in Figure 14. The top picture shows the original, the middle one the target/background segmentation on full frames and the bottom one, the extracted components of a tank target.

2.2.2. Feature Extraction

The image over an extracted candidate object is processed to extract the following features:

- Total length, L

- Total area, A

- length-to-width ratio,

- Area-to-perimeter squared,

- Average contrast

- Moments: a_{00} - a_{30}, a_{03}

where

$$a_{pq} = \iint_{object} x^p y^q I(x,y) \, dx \, dy$$

where $I(x,y)$: intensity at (x,y)

The features are used by the classifier for training and testing the classifier.

2.2.3 Classification

Target classification for a small image, i.e. when the number of pixels occupied by the target is small, can best be done by statistical methods. However, at short ranges, or high magnification, "syntactic" classifier is used to exploit the available structural information.

2.2.3.1 Statistical Classification

Statistical classifiers are rules for partitioning the feature space such that target classes are separated from each other and from non-target objects by the partitioning surfaces. Design of a statistical classifier requires a knowledge of the target and non-target feature statistical distributions. These are usually provided in the form of feature samples, extracted from representative imagery. Determining the optimum partitioning from the sample set is called "training" the classifier. A separate, statistically independent feature data set is used for "testing", or estimating the performance of the classifier.

Three types of classifiers that we use are linear, discriminant tree and k-nearest neighbor classifiers. These classifiers can be used for either clutter rejection or discriminating between classes of targets.

There are several approaches to the design of a linear discriminant classifier. The most common utilize data from two classes at a time. If the data is bimodal and linearly separable a simple and effective approach is to use a Fisher linear classifier. If we define:

$$W = (C_1 + C_2)^{-1} (M_1 - M_2) \qquad (9)$$

$$T = W^t (\sigma_1^2 M_2 + \sigma_2^2 M_1)/(\sigma_1^2 + \sigma_2^2) \qquad (10)$$

where C_1, C_2, M_1, M_2 are the covariance matrices and means of the two sets of data and $\sigma_1^2 = W^t C_1 W$, the Fisher linear classifier is given by:

$W^t X > T$ implies X belongs to class 1

$W^t X \leq T$ implies X belongs to class 2

where X is the feature vector of the object being classified, and T is the optimum threshold.

A tree classifier is simply a logic tree in which each branch consists of a linear discriminant. In its simplest form, it consists of a sequence of cascaded thresholds on the features. These two approaches work well when only two classes are to be discriminated and when the feature vectors are linearly separable. For a more general case involving more than two classes and the feature vectors not separable, the k-nearest neighbor (kNN) classifier is more suitable. In the kNN approach, the set of training and the class types of the samples are stored. A new sample vector X is classified as follows. Its distances to all the stored training samples are found, and the k nearest are found. The new sample is assigned to the class belonging to the majority of its k nearest neighbors.

Using the kNN classifier on FLIR images, we have been able to achieve nearly 80% correct classification between two types of military targets. Better performance can be expected in some scenarios by further processing of frame-to-frame sequences of decisions. These smoothed decisions reduce the effects of random errors in the individual decisions.

Target screener performance can also be improved by using range data, measured either directly or indirectly. Range makes possible absolute size discrimination which is lacking in FLIR imagery.

2.2.3.2 Syntactic Classifier

There are cases where, the targets show visually distinguishable shape and internal intensity variation. Statistical classifications of this type of targets would require an unmanageably large number of features and complicated statistical distribution. In picture recognition problems, when the number of features required is very large, the concept of describing complex patterns in terms of a (hierarchical) composition of simpler subpatterns becomes very attractive (Aggarwal, R.K., 1978). The description of a target in terms of simpler subpatterns, i.e., target components, enables us to perform syntactic recognition of targets.

The assumptions in this approach to tactical target recognition are:

- Images of tactical targets are "large" enough to show structures.

- It is easier to recognize target components than the target.

The first assumption deals with the sensor-target range. If the range is too long to show any details inside the target image, one would have to resort to statistical recognition techniques. But as the sensor-target range decreases and the target structure becomes discernible, syntactic recognition schemes become feasible.

The second assumption deals with the relative ease of recognizing the target and its components. If it is easier to recognize a target than its components, as would be the case when the target image is only a few pixels, one would not employ syntactic recognition schemes. But in low-quality images where the recognition based on target outline is not very reliable, a syntactic scheme can be successfully used to recognize targets, provided the assumption on target image size holds. Syntactic recognition schemes can also be successfully used for partially occluded targets where conceivable statistical recognition schemes would fail.

Consider, for example, the image of a tank depicted in Figure 15. Suppose it is possible to recognize the components of this tank, such as motor, hot spots, vents, barrel, etc., using statistical properties of each component and their spatial relationship. A hierarchical (tree-like) structural information in this tank can be represented by a tree as shown in Figure 16. Syntax rules can be used to describe tree structures like the one in Figure 16. The syntax or grammatical rules for this example are:

TANK → RECTANGLE, HOTSPOTS, BARREL

RECTANGLE → TREAD, MOTOR, VENTS

Since different components of a target may be seen from different aspect angles, a general set of rules can be inferred by training the classifier with tree structures of the target viewed from different aspect angles. The general block diagram of syntactic approach to tactical target recognition is shown in Figure 17.

Syntactic tank-recognition scheme is discussed below as shown in Figure 18. The feature set of each segmented component is stored in a "label table" comprising the description of the component. Classification of these segments to individual candidate components is performed before syntactic recognition can be completed. This component classification is done applying statistical technique to the feature vector in the label table.

The goal of syntactic tank-recognition algorithm is to use the label table for a possible tank and decide whether the object can be classified as a tank, based on identification of the components. First the motor is identified as the initial classification step. If no additional tank components can be further identified, boundary shape analysis is required for recognition. An approximate orientation of the object is determined using the already defined body and motor groups. Direction of the possible tank is determined by examining the location of the motor group relative to the body group. Search regions are established for locating hot spots, vents, and barrel of the tank. Size, shape, and direction features are used for component recognition. If at least one additional tank component is found, the object is declared a tank. Otherwise, statistical methods based on object boundary features are needed for further classification.

Honeywell has extensively tested syntactic target recognition technique on FLIR tactical targets. Figure 19 shows an example of syntactic tank recognition in FLIR image.

2.2.4 Symbol Generation

Once a target is detected and classified a symbol is generated and displayed in (x,y) position. There is a symbol for each class of targets and its size is proportional to the size of the target image. The symbol may be updated every time a decision is made.

3. CONCLUSION

This paper presented a concept for an autonomous target screener. Based on the material presented, the following conclusions are made:

- The feasibility of detecting and recognizing tactical targets in thermal IR imagery has been proven over a limited data base.

- An autonomous target screener will require both statistical and syntactic classifiers.

- An integrated FLIR, Image Enhancement, Target Screener system is possible in the near future.

REFERENCES

1. KREBS, M.J., and LORENCE, L., October 1974, "Real-Time Display Parameters Study: II, FLIR Target Acquisition as a Function of Training and Imagery Characteristics", Final Report Contract F30602-73-C-0377.

2. NARENDRA, P.M., and FITCH, R.C., May 1978, "Real Time Adaptive Contrast Enhancement for Imaging Sensors", SPIE, Volume 137.

3. NARENDRA, P.M., et al., December 1977, "Final Report on Automated Image Enhancement Techniques for Second Generation FLIR", Honeywell, Contract No., DAAG53-76-C-0195.

4. PANDA, D.P., SERREYN, D.V., GEOKEZAS, M., March 1978, "Autoscreener with Secondary Screener", SPIE Volume 137.

5. STANO AIRBORNE SYSTEM Part II Test Report (U), July 1971, U.S. Army MASSTER, Ft. Hood, Texas.

6. AGGARWAL, R.K., 1978, "Adaptive Image Segmentation Using Prototype Similarity", Proceedings of IPPR Conference, May 1978, IEEE.

7. AGGARWAL, R.K., November 1978, and Wittenburg, T.M., "Syntactic Recognition of Tactical Targets", Proc. Image Understanding Workshop, Nov. 1978, DARPA.
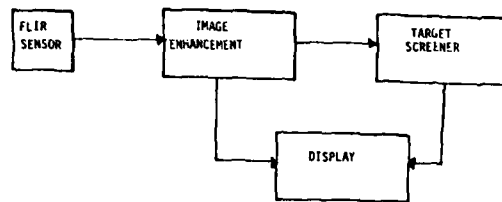
Figure 1.   Autonomous Target Screener.
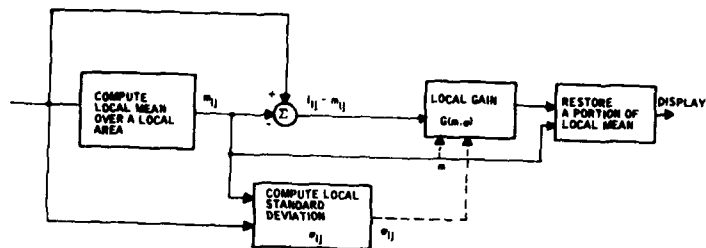
Figure 2.   Functional Block Diagram of the
Adaptive Contrast Enhancement Algorithm.
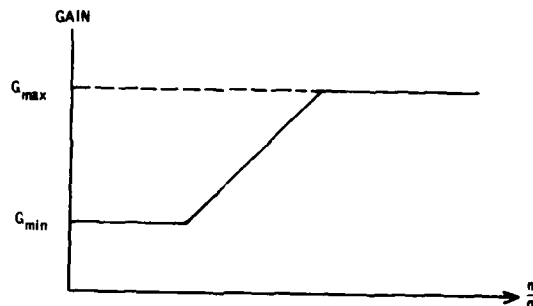
Figure 3.   Local Area Gain Curve to Prevent
Excessive Gain Variations.
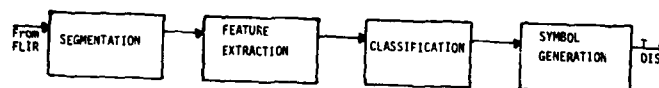
Figure 4.   Target Screener.
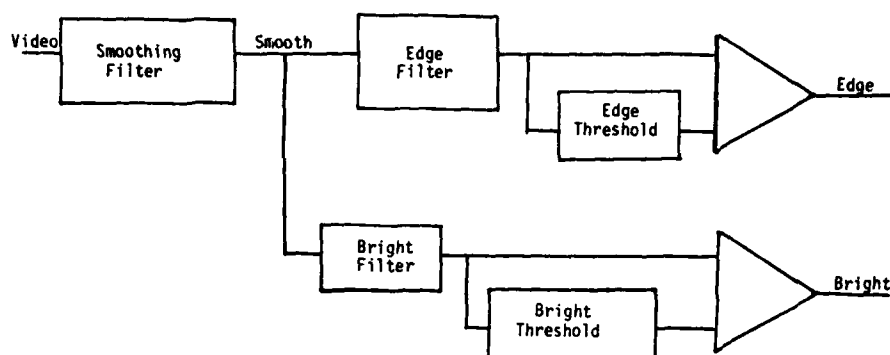
Figure 5. Autothreshold Functional Block Diagram.
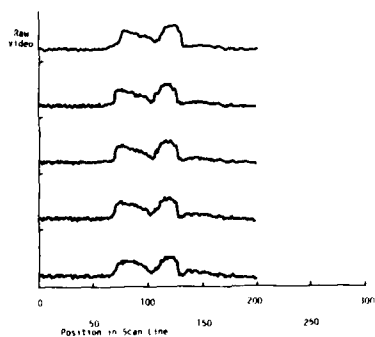
Figure 6. Five Scan Lines of Raw Video Over a Tank.
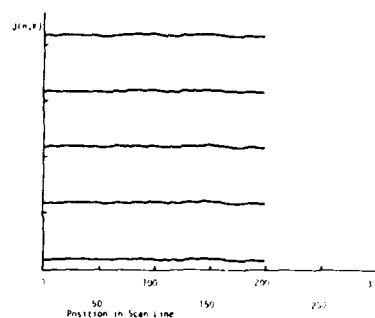
Figure 9. Background Estimate of Data in Figure 4.
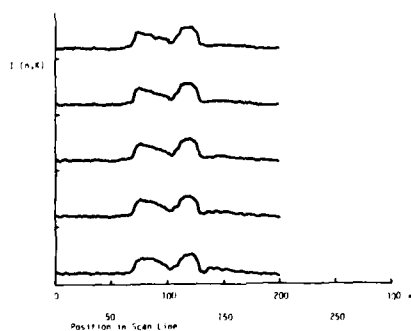
Figure 7. Smooth Video of Data in Figure 3.
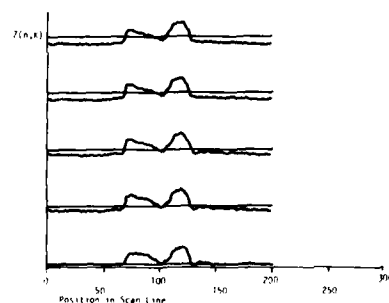
Figure 10. "Video-Background" of Data in Figure 4.
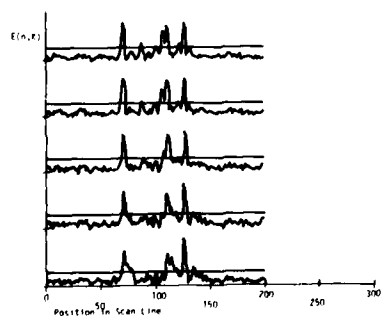
Figure 8. Edge with Threshold of Data in Figure 4.

11a. FLIR Image

11b. Binary Edge Image
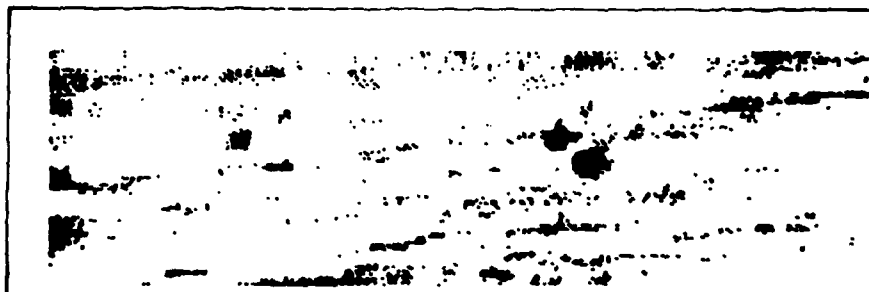
11c. "HOT" Thresholded Binary Image

Figure 11. Image Segmentation
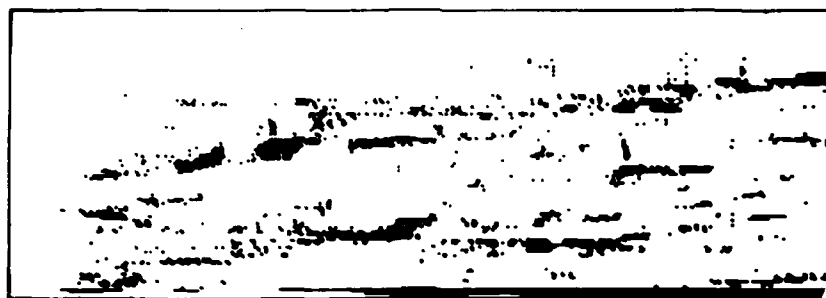
11d. "COLD" Thresholded Binary Image
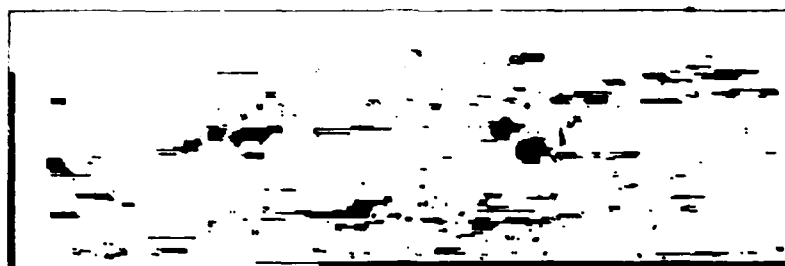
11e. "OBJECT INTERVALS"
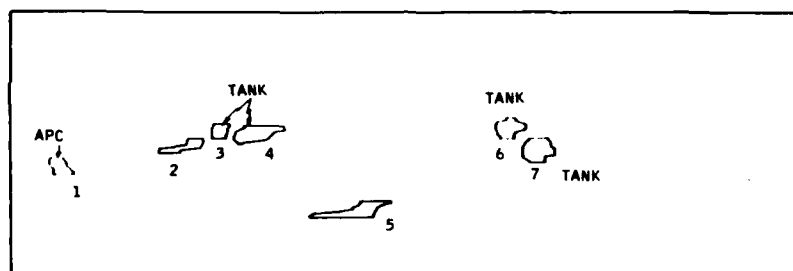
11f. Extracted Objects

Figure 11 (cont.). Image Segmentation.
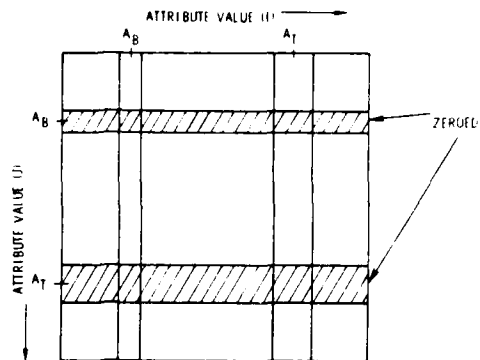
Figure 12. Two-dimensional Distribution
Function f'(j,i).
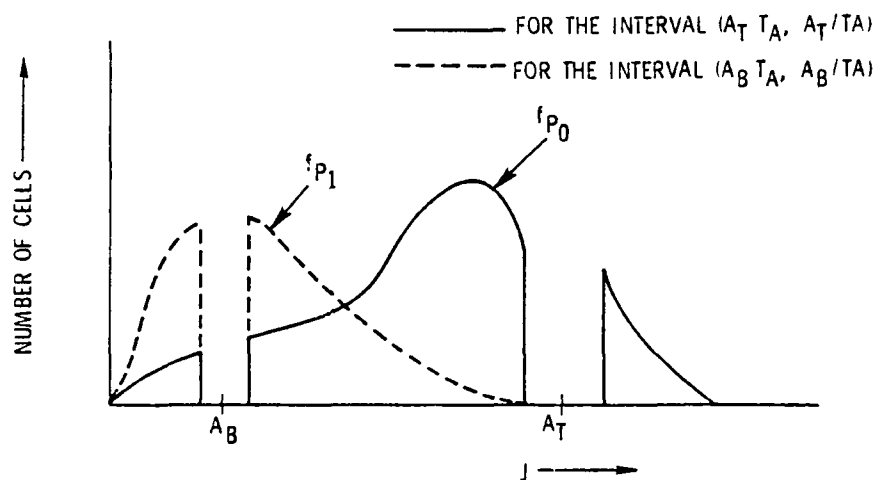
Figure 13. Accumulative Attribute Profiles
f_{P_0} and f_{P_1}.

Figure 14. (a) FLIR Image of a Tank, (b) Tank Detection at low
Resolution, (c) Tank Detection at High Resolution.
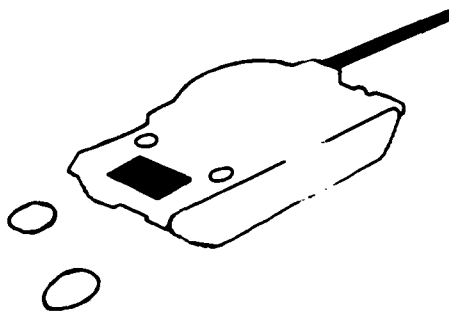
Figure 15.   Image of a Tank.
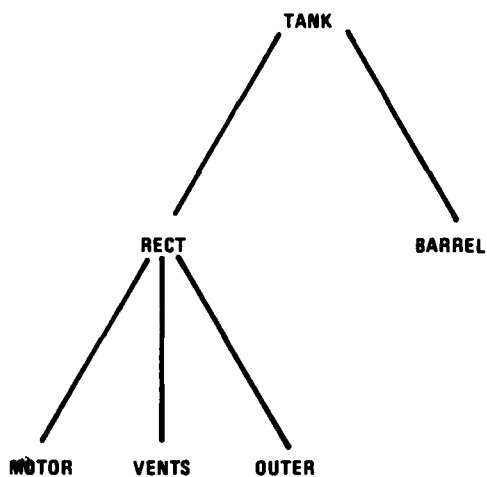
Figure 16.   Hierarchical Structural Description
of the Tank in Figure T5.
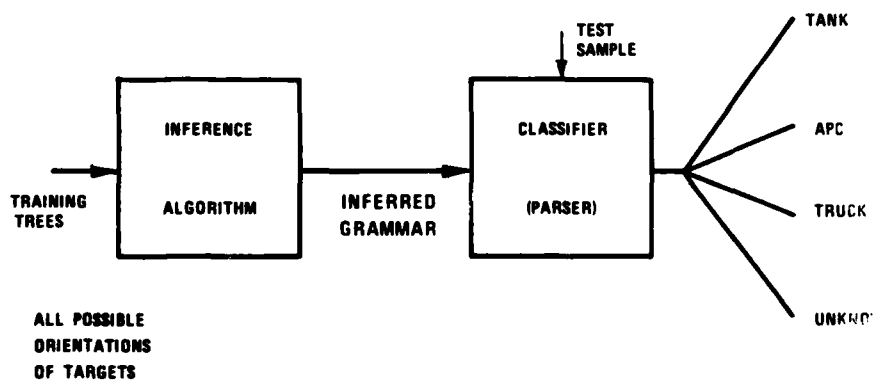
Figure 17.   Syntactic Approach for Tactical Target Recognition.

Figure 18. Syntactic Tank Recognition.

SYNTACTIC RECOGNITION

INPUT IMAGE

COARSE SEGMENTATION

COMPONENT EXTRACTION

COMPONENT CLASSIFICATION
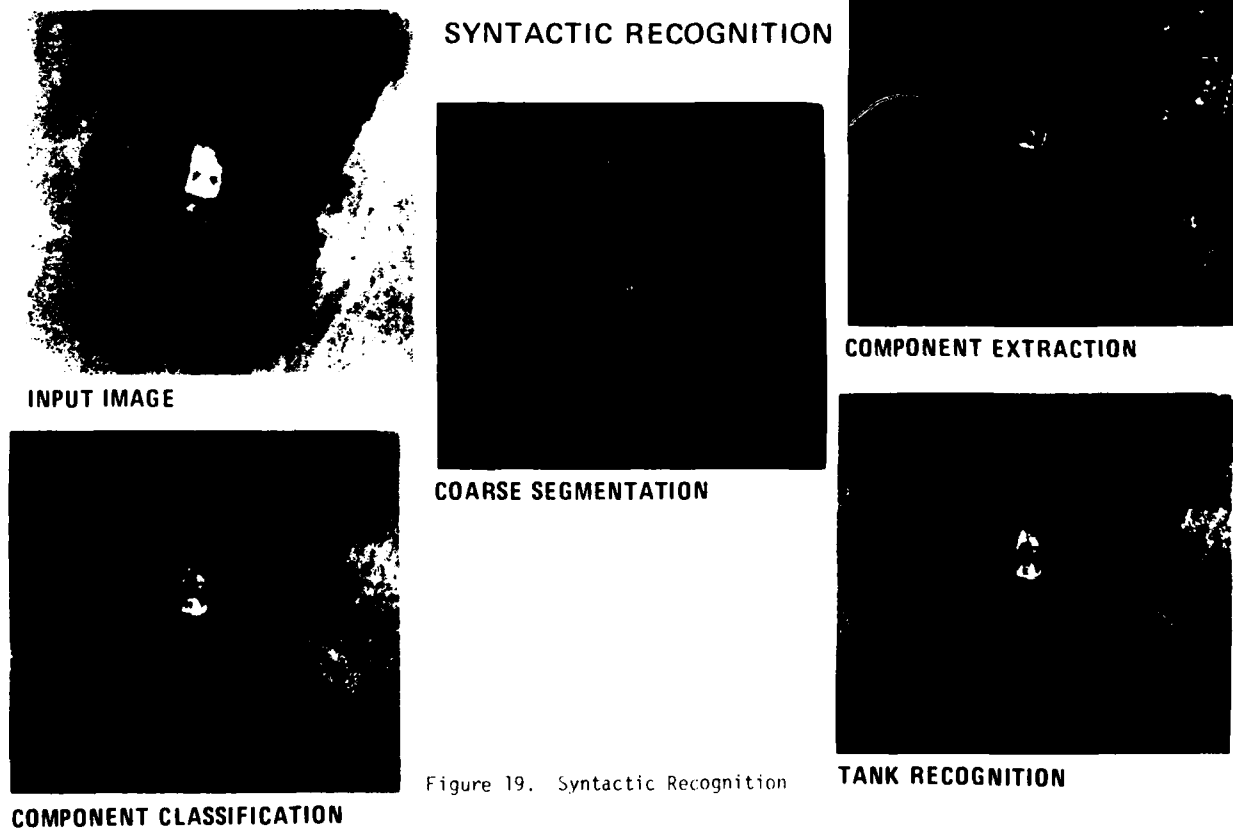
TANK RECOGNITION

Figure 19. Syntactic Recognition

EVALUATION OF A REALTIME PROTOTYPE AUTOMATIC TARGET CUER

Terry L. Jones
U.S. Army Night Vision & Electro-Optics Laboratory
Fort Belvoir, Virginia U.S.A.

SUMMARY

The procedures used and problems encountered in the preliminary evaluation of a proto-
type automatic target cuer which operates on video imagery are presented. Performance
measures are defined and their use and limitations described. An overview is also
given of the operation, training, and field test scenario used to test this target cuer.
Results of cuer performance are presented as examples of cuer evaluation procedures.

1. INTRODUCTION

In March of this year Honeywell Systems and Research Center delivered to our Laboratory
the prototype automatic target screener (PATS) for preliminary flight testing. PATS
incorporates many of the algorithms discussed in the previous paper (Aggarwal, 1980),
and some details will be provided here. However, this paper primarily describes the
evaluation of automatic cuer performance. The main objective is to provide insight into
the complexity of this task. By stimulating thinking on this task now, it is hoped that
evaluation procedures will be established which will allow the NATO forces to ask for
and get the image processing capability needed in the future. This paper hopes to
achieve this goal by presenting examples from the preliminary testing of PATS.

2. PROTOTYPE AUTOMATIC TARGET SCREENER (PATS)

The PATS hardware is shown in figure 1. The main unit is 60cm long, 20cm high, and
25cm wide. With the small control box, shock mounts, and cabling, its mass is 28kg.
It requires about 200W for operation. It can operate on either 525 or 875 line, 2:1
interlaced, 60 field per second video. All operations are performed on a single video
field. When processing of that field is completed, another field is processed. Because
of this, the processing time beyond the preprocessor stage is image dependent; PATS
occasionally processes consecutive odd or even fields, but typically processes only
every sixth field.

The functions performed in PATS are shown in figure 2. The preprocessing section is
primarily used to provide maximum information transfer from the sensor to the processor
by controlling global gain and bias settings of the sensor. However, operation with
an AC coupled detector array imposes a need for DC restoration, so this is also in-
cluded. In addition, the preprocessing section performs local area gain and brightness
control (LAGBC) for improving the imagery displayed to the operator.

After preprocessing, the image is segmented into objects of interest. The first step
in this process is the formation of intervals of interest on individual scan lines.
These intervals are formed by the matching of adaptively thresholded bright and edge
signals for up to about twenty intervals per line. The resulting intervals are corre-
lated line to line in bins which are filtered to form the objects of interest.

Next, simple features are calculated for the objects of interest. This process actually
begins during the field scan, and these features (such as interval position and bright
count) are stored in the data memory. The features are used to discriminate between
clutter objects and suspected targets by a series of tests comprising the clutter re-
jection classifier stage.

More complex features (such as high order intensity moments) are then calculated on the
remaining objects. At this stage there can be no more than about twenty objects, or
processing time becomes too long to do object tracking in dynamic imagery. The complex
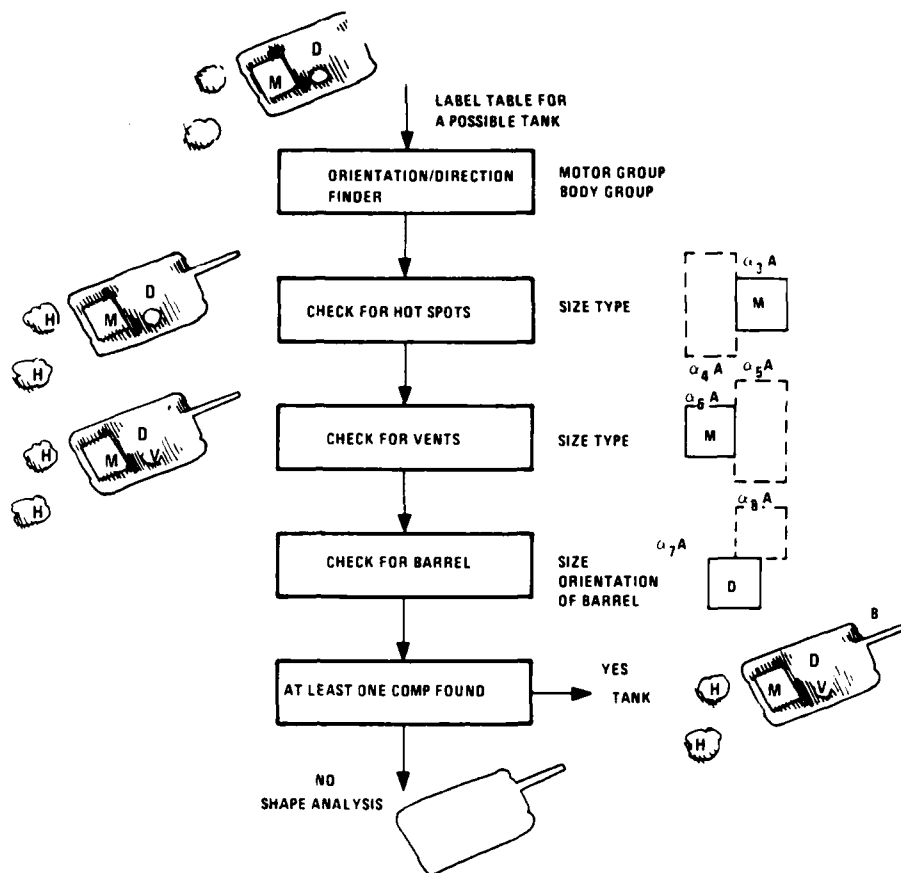features of these objects are then compared to features of prototypes of up to five
classes that have been stored in the memory. Each object is classified based on a
majority of nearest neighbor prototypes. The object classes are easily changed by
replacing the features, which reside in eraseable, programmable, read only memory
(EPROM), with new prototype features. The prototype capacity is limited to about
1000 for realtime operation.

After the classification decisions are reached on all objects in the field being pro-
cessed, the results are stored, and a new field is processed through the same procedure
generating another set of classified objects. The interframe analysis section does a
processed field to processed field correlation of the classified objects and keeps
track of the string of classification decisions made on each object. A multi-frame
classification decision for each object is then made based on the last five to fifteen
single field decisions. A symbol for each object is then generated based on the multi-
frame classification and added to the video at the location of the last segmentation of
each object classified.

3. PATS TRAINING

The PATS hardware can operate on standard video from any sensor. However, PATS must
first be trained on the sensor's output. The sensor chosen to provide PATS' first

training imagery was the common module, thermal imaging system called LOHTADS (Light Observation Helicopter Mounted Target Acquisition/Designation System). Since this system can provide range information to the center of the field of view, PATS' clutter rejection stage was trained to use this information in addition to the intensity information. A limited computer simulation was done by Honeywell on digitized images to select the initial features to be incorporated in the software. Once the hardware was fabricated, it was used to do this process from analog videotapes. The edge and bright thresholds for generating intervals are adjustable in the hardware and are set to optimize the segmentation. The clutter rejection stage is programmable with the microcode residing in EPROM. Although the microcode can be changed in a matter of minutes, large changes require a time consuming checkout, and even a simple change, like adjusting the value of a limit for an area test, requires several hours of analysis from different videotape inputs to assess its overall impact on clutter rejection performance. Training the classifier requires only a matter of hours once the prototypes have been selected. But once again, choosing the prototypes which give a representative sample of the existing data is a difficult and time consuming task. The initial classes chosen for PATS were M60 tank, M113 armored personnel carrier, M35 flatbed 2½ ton truck, and M151 jeep. Videotapes from LOHTADS of these four classes in all aspects and with engines running were used to train the classifier. Since PATS is a statistical classifier, all training and testing was done on targets with no partial obscuration. To date, all training tapes have been generated at Fort A.P. Hill, Virginia.

4. FLIGHT TEST SCENARIO

PATS was mounted in the rear compartment of a light observation helicopter (OH-6) in March for preliminary field testing. Along side PATS were two videotape recorders, one to record the input to PATS, the other to record the PATS output. The test was performed at Fort A.P. Hill, Virginia. Most of the fort is hilly (less than 100m) and forested with a mixture of deciduous and coniferous trees. The target vehicles were always located in the drop zone, a cleared area about 1km by 4km. Most of the drop zone is covered with small shrubs, but there are bare spots and several clumps of small trees. There are portions of paved and gravel roads through the zone, but more prevalent are the many dirt trails. For this test, most of the vehicles were stationary; however, the vehicles were always exercised for at least 20 minutes prior to collecting data and their engines left running during the test. This was done because all training data was of hot vehicle. Videotapes were generated as the helicopter flew from several kilometers away to within 0.5km of the targets. The altitude of the helicopter never exceeded 200m above ground level. The following data were collected during the test:

    a. 875 line videotapes of both LOHTADS and PATS output with range and time on audio channels.

    b. Weather data (temperature, humidity, wind velocity, barometric pressure and cloud cover).

    c. Aircraft heading and altitude.

    d. Target vehicle information (location, orientation, and radiant temperature).

The initial test plan was to reproduce as closely as possible the imagery used in training PATS. It became clear the very first day that this task could not be achieved to our satisfaction because major differences between the two image sets were being caused by different environmental conditions.

5. PERFORMANCE MEASURES

With the appearance of automatic image processing equipment for military applications, it becomes necessary to define quantifiable performance measures that can be used to assess approaches and equipment. It is important to decide if a cuer is good enough for a specific task, to choose between one processor and another, to compare men to machines, and to measure development progress on specific hardware and progress of the technology in general. In this section the techniques used to describe cuer performance are presented. The actual parameters used in a general performance measure are defined for both detection and recognition tasks. This measure is then used with examples from PATS testing to describe absolute cuer performance on a given data set, to make comparisons with human observers, and to measure cuer improvements. A proposed method of comparing cuers is discussed in the last portion of this section.

5.1 PARAMETERS USED TO EVALUATE PERFORMANCE

Most measures of performance are defined for a specific mission and are only useful when trying to evaluate an automated system for that particular application. For example, an effectiveness, E, can be calculated for a system on a given set of images of size X by Y using

$$E = \sum_{x=1}^{X} \sum_{y=1}^{Y} \sum_{i=1}^{n+1} \sum_{j=1}^{n+1} A_{ijxy} P(i,j) P_{jxy}$$

where P_{jxy} is the probability of an object appearing which belongs to class j at location xy; $P(i,j)$ is the conditional probability that if an object belonging to class j appears, that object is called a member of class i; and n is the number of classes of interest. (The i and j are summed to n+1 to include the class of no interest objects.) The coefficients in the sum are positive for correct decisions (i=j) and negative for incorrect decisions (i≠j) with the magnitudes assigned based on the relative importance of each decision for the particular mission. (For objects of no interest, $P(n+1,n+1)$ does not need to be measured because $A_{x,y,n+1,n+1}=0$). The best overall performance is achieved when E is maximized.

Other possible measures can be defined which are less mission dependent than E. For example, the probability of not detecting a target in class j (assigning it to the non-target class) is simply $P(n+1,j)$. The probability of target recognition, P_R, is the number of targets correctly classified divided by the total number of targets present. The probability of incorrect target classification, P_I, is just the number of targets incorrectly classified divided by the total number of targets present. These are more general measures than E because they combine many of the terms used in E. Therefore, they cannot be used to predict the performance of a system in a particular scenario as well as E can. However, they are more useful than E at this stage of cuer development because they can be more easily used to describe general performance and to make comparisons.

Unfortunately, no one of these general measures is very useful by itself because rarely will only one change when the cuer or its algorithm changes. For example, increasing P_R would generally be considered good, but even doubling the value of P_R is probably bad if doing so also means that P_I is tripled. Reducing P_R by 10% is probably very worthwhile if P_I can be cut in half. The ratio of P_R to P_I is a better general performance measure than either alone; but this is also inadequate because it does not include a measure of the difficulty of the task being performed on the image set. For example, the task of distinguishing between tracked vehicles and wheeled vehicles (a two class task) is an easier one than the task of deciding among M60 tank, M48 tank, M551 light tank, M113 APC, M35 truck, and M151 jeep (a six class task) in the same image set. Therefore any performance measure should be a function of n, the number of classes. In addition, it is important to note that any image set has a variety of difficulty of the recognition tasks within it for even a fixed n; classifying 100% of the targets is more difficult than classifying 10%, while refusing to decide on the classification of the other 90%. This difficulty is at least partially proportional to P_R. There is not a simple yet satisfying way to take these two task difficulties into account, but one way is to multiply the P_R/P_I ratio by $n^2 P_R$. One further item that must be included in the measure is the time allowed to make the classifications. If performing the operations twice as fast is considered about twice as good, the inverse of the average processing time can be simply included as a multiplicative factor. Taking the above considerations into account, the following recognition performance measure was chosen for the initial evaluation of PATS:

$$pM = \frac{(nP_R)^2}{P_I t},$$

where t is the time interval upon which classification decisions are based. (In PATS testing t is the number of fields used in the interframe analysis times the average time between field grabs.) This pM has a range from zero (bad) to infinity (perfect). Table 1 shows the range of values expected from cuers over the next few years along with subjective ratings.

Table 1: Illustration of the range of pM

Subjective Rating	Number of Classes, n	Probability of Recognition, P_R	Prob. of Incorr. Classification, P_I	Time Required for Decision, t	Performance Measure, pM
bad	2	0.2	0.2	2.0s	$0.1s^{-1}$
fair	3	0.3	0.1	1.0	0.9
good	4	0.4	0.05	0.5	6.4
excellent	5	0.5	0.025	0.25	42

If the above pM is to be used to describe detection, then P_R must be replaced by the probability of target detection, P_D, the number of targets detected divided by the total number of targets. But in this case there is no longer a need to include a measure of the difficulty of the task the cuer is doing on a given data set because the task is fixed: the cuer must always choose between target and non-target for each position on every image. Therefore, the numerator need only consist of $4P_D$. ($n^2=4$ is retained to make the subjective ratings in table 1 still appropriate when discussing detection). P_I, however, must be replaced by a measure of both targets called non-targets (missed targets) and non-targets called targets (false alarms). Since this measure is only to be used over a particular data set, it is not necessary to define a false alarm rate. P_I can be replaced by the sum of the probability of missing a target, $P_m=(1-P_D)$, and the number of false alarms per target, F. The time is inserted in the same manner

as before yielding a performance measure for detection,

$$pM_d = \frac{4P_D}{(P_m+F)t} .$$

The performance measures just described are probably not the best to use; there are several problems with them. The measures could become unrealistically large as the denominators go to zero. However, this occurrence probably indicates a lack of an appropriate number of measured samples as much as a fault with the pM. Another problem is that when n is larger than two, the classes are never going to be equally easy to separate. Despite these and other reservations, these measures are the best we have at the moment so they are being used to evaluate cuer performance.

5.2 ABSOLUTE PERFORMANCE MEASURES

The general pM's described above can be used to rate the performance of a cuer on any given set of data. Table 2 shows the results of such measurements for five data sets collected for similar scenarios with four targets located in the center of the drop zone at Fort A.P. Hill, Virginia. Since the cuer operation was not different between these scenarios, all that the pM values show is the differences between data sets used in testing. To establish an absolute value for any performance measure that can be generalized to the real world performance, the distribution of samples in the data set must accurately reflect the variety in the real world. Acquiring sufficient thermal imagery to form this data base is clearly impractical. Simply specifying what should be included is beyond the state of the art in image description.

Table 2: Absolute cuer performance (PATS)

Date	Time	t	P_R	P_I	pM	P_D	F	pM_d
15 Nov 79	1530 hrs.	1.5s	.16	.30	$0.9s^{-1}$	.57	.01	$3.5s^{-1}$
12 Mar 80	1530 hrs.	4	.05	.06	.15	.44	.26	.5
14 Mar 80	1630 hrs.	5	.03	.10	.02	.58	.82	.4
15 Mar 80	1600 hrs.	5	.04	.03	.17	.45	.38	.4
15 Mar 80	2000 hrs.	6	.07	.15	.09	.56	.64	.3

An alternative to forming this massive data base is the use of theoretical models (verified over a limited data base) to predict the full spectrum of variety based on measurements other than thermal imagery. Target orientation, target temperature, time, season, solar insolation, humidity, and other parameters would be used with this model to predict the "inherent-detectability or recognizability" of a target in a given thermal image set; then the results of an absolute pM on a cuer could be compared to the "inherent characteristic" and the cuer performance rated. Unfortunately, the current status of thermal image modeling is not sufficient to predict this "inherent characteristic" from a set of measured parameters. This situation is not expected to change in the near future so it is important that another alternative be formulated to compare cuer absolute performance on different data bases.

5.3 COMPARISON WITH OBSERVERS

One way to gauge the difficulty of the detection or recognition tasks assigned to cuers on specific data sets is to present the imagery to an average human observer. The results of the observer's performance can then be compared to the cuer's. This procedure is filled with pitfalls, but to date it has been the most reliable method of assessing cuer performance. The average observer, of course, is a myth. Instead, a group of observers is used to smooth out the person to person and day to day variation of individuals. This group must also be checked periodically on the same data to determine group drift. In addition, it is important that tests given to the observers do not allow them to learn during the testing. Since this is next to impossible, it is important to design the tests so that results from observers who do learn during the testing can be eliminated from the group. For example, one set of imagery shown to the observers for measuring observer pM contained target vehicles that were always oriented in the same direction as each other within each image. One observer discovered this during testing. Knowing the orientation of all vehicles, once the orientation of one vehicle could be determined, allowed this observer to perform better than if he had not learned this information. The primary method used to determine observer learning is to repeat the first portion of the test as if it is simply a continuation of the original test. If an observer improves his performance on that portion the second time, his results must be discarded.

Handling the time an observer is allowed to view and make decisions on the test imagery is also an issue that must be decided carefully. For most of the comparisons that have been made to date, the observers viewed only a few seconds of videotape and were forced to respond during the viewing in an effort to force a response time close to PATS'. However, another test has been set up to task load the observers during the viewing

of thirty seconds of tape. This more nearly simulates the field environment but is difficult to use for a direct comparison with the cuer. It is probably more useful in comparing operator performance with and without a cuer, something we hope to do shortly.

Given the above limitations, tables 3 and 4 show typical results of these types of comparisons for representative scenarios during the preliminary flight test of PATS in March. All the comparisons shown use night imagery because the daytime imagery that would be useful for this comparison was too cluttered for either the observers or PATS to perform at accurately measurable levels.

Table 3: Recognition performance of observers compared to PATS

Date	Time	Obs. n	Obs. t	Obs. P_R	Obs. P_I	Obs. pM	PATS n	PATS t	PATS P_R	PATS P_I	PATS pM	$\dfrac{\text{Obs. pM}}{\text{PATS pM}}$
14 Mar 80	2230 hrs	4	5s	.21	.42	$0.3s^{-1}$	4	5s	.06	.06	$0.2s^{-1}$	1.7
15 Mar 80*	1945 hrs	4	5	.49	.22	3.5*	4	5	.07	.15	0.11	32*

*For this test the observers' task differed from PATS' because the observers knew before their five second observation period the target array pattern and the array's location.

Table 4: Detection performance of observers compared to PATS

Date	Time	Obs. t	Obs. P_D	Obs. F	Obs. pM_d	PATS t	PATS P_D	PATS F	PATS pM_d	$\dfrac{\text{Obs. } pM_d}{\text{PATS } pM_d}$
14 Mar 80	2230 hrs	5s	.89	.05	$4.5s^{-1}$	5s	.46	.34	$.42s^{-1}$	11
15 Mar 80	2130 hrs	5	.85	.11	2.6	5	.15	1.5	.05	51

5.4 MEASURES OF CUER IMPROVEMENT

Like all newly emerging technologies, cuer performance will improve rapidly with time. It therefore is necessary to measure this improvement. Because of the reasons described in section 5.2, no absolute measure of performance can be used for this task when trying to compare results from two separate field tests. There are two avenues which can be pursued. First, one could compare cuer operation at two different times on the identical input data. It is important when trying to assess general improvement to use different test data than training data. This is always impossible, but one strives to do the best he can with the available data. This was done with PATS in July 1980 by inputing video-tapes from November 1979 and March 1980 with PATS on the bench. It was also necessary before drawing conclusions from this test to show that PATS operated the same on the bench as it did in the aircraft. After having shown this, PATS' detection performance was compared before and after some major hardware and software changes. The ratio of July to March performance showed an improvement by a factor of three. This was considered significant enough to attempt another flight test in August 1980. Results of that test will be forthcoming. In early August, just prior to the flight test, PATS was tested on a tape from which no training samples had been taken. As shown in table 5, PATS detection improved by a factor of five.

Table 5: Measure of PATS improvement

Date of Test	t	P_D	F	pM_d
1 Aug 80	3s	.79	.37	$1.8s^{-1}$
15 Mar 80	6	.56	.64	.35

A second method of measuring cuer improvement can be used if one has a measure of the difficulty of the task. As described in section 5.3, observers can be used to provide this standard. How well the cuer had done compared to an average observer on one data set can then be compared to how well the improved cuer does compared to an average observer on the new data set. This method will be used to report the results of the August tests as they become available.

5.5 COMPARISON BETWEEN CUERS

As other realtime cuers are developed and field tested, it will become essential to compare the strengths and weaknesses of each. The ideal procedure would be to compare

them on the same input data, but this will frequently be impossible. For example, Northrop is fabricating a cuer which will be ready for preliminary testing next year. This cuer is also designed to operate on LOHTADS thermal imagery, but it incorporates a digital scan converter in the front end. It cannot operate on the video formated imagery required for PATS operation. It might be possible to modify the Northrop cuer just to furnish a video output that could be fed to PATS, but even this comparison would not be perfectly valid because PATS would be forced to operate on video from an imaging system whose gain and bias it is not controlling. Most comparisons between cuers will not even be as simple as this one since they will not be able to operate from the same sensor.

One way to make the comparison is through the average observer. The observer gives a measure of the difficulty of the task, and cuers are compared by measuring how much better than observers they do. Once again, the more similar the test data, the more confidence there will be in the result. If cuers evolve to do different tasks, this comparison will never be made properly because the two cuers will not be trained with similar data.

6.0 CONCLUSIONS

Procedures are being developed and used to evaluate realtime target cuers. These procedures need to be greatly refined and improved in order to meet the rapidly growing need for quantitative cuer assessment.

7.0 ACKNOWLEDGEMENTS

The author wishes to acknowledge the many useful discussions with colleagues at both the Night Vision and Electro-Optics Laboratory and personnel at Honeywell SRC. Joe Swistak of NV&EOL has been a major contributor in this effort by formulating PATS test procedures. The Allen Corporation has done most of the PATS data reduction.
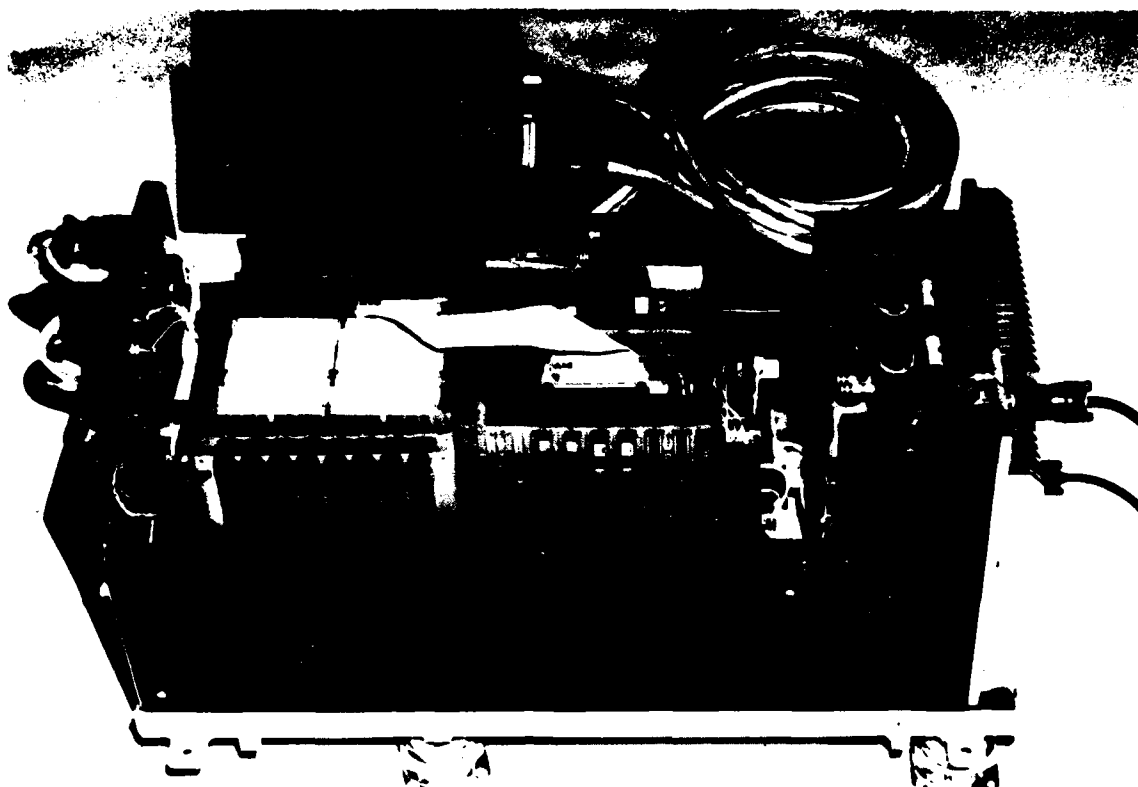
FIGURE 1: PATS HARDWARE WITH THE TOP COVER REMOVED.
THE SMALL CONTROL BOX IS MOUNTED BETWEEN
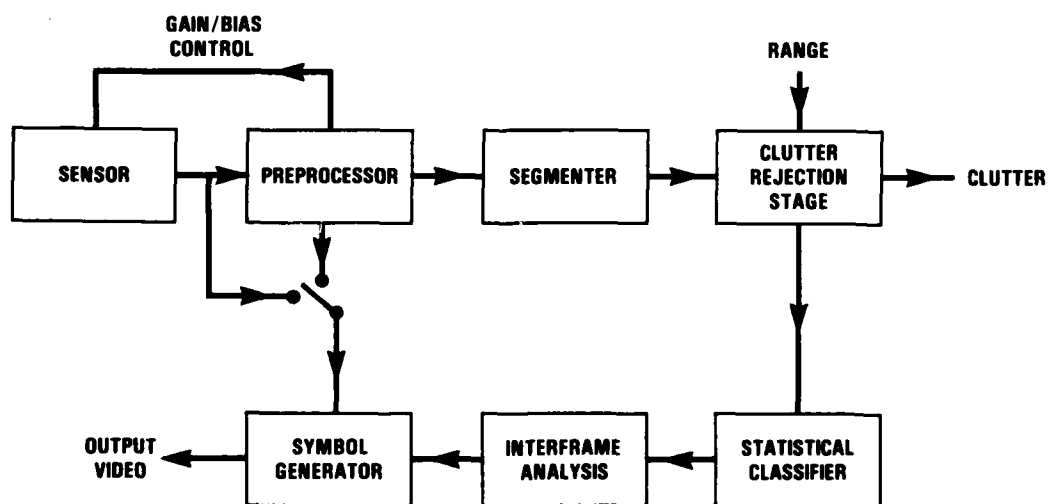THE TWO SEATS IN THE HELICOPTER.

FIGURE 2: BLOCK DIAGRAM OF PATS' OPERATION.

DIGITAL IMAGE PROCESSING FOR GROUND TARGET DETECTION, IDENTIFICATION AND LOCATION

George R. Hughes

RADC/IRR
Griffiss AFB, New York

ABSTRACT/SUMMARY

This paper discusses RADC/USAF Research and Development activities in the areas of image processing and target identification relating identification and location of ground targets in a tactical scenario. The primary objective is to reduce the reconnaissance cycle from days/hours to minutes and seconds commensurate with the near real time (NRT) intelligence requirements of the tactical forces. This NRT intelligence is essential to support strike of increasingly mobile enemy weapon systems. In currently fielded systems exploitation of film based reconnaissance is extremely slow, greatly lagging collection rates.

There are three essential elements to a NRT tactical intelligence system. They are NRT imagery collection, NRT air to ground image data link and NRT imagery exploitation. The thrust of this paper will be to discuss technology required to support the later, NRT imagery exploitation.

Technology intensive efforts are categorized under each of the exploitation elements (target detection, identification and precision location). R&D in the area of target detection consists of exploratory and advanced development work units in automated target correlation, automatic change detection and pipeline image processing for screening probable target areas. Target identification R&D to be presented includes automatic techniques for pattern recognition as well as semi-automated techniques for aiding an analyst by correlating various sensor and intelligence inputs to permit target identification. Near real time precision target location techniques will include techniques for locating imagery targets in a predefined precision photographic data base as well as techniques for performing location simultaneously with target identification.

Summary and conclusions for this presentation discuss commonality aspects of a digital image exploitation system relating to a "universal" image exploitation system and the potential for its use in NATO. Potential areas for developing cooperative R&D programs in support of this universal system objective will be identified.

BACKGROUND

A primary mission of tactical Air Force is to neutralize enemy ground combat forces thru strike of second echelon ground support forces. Optical reconnaissance sensors, using hardcopy film as a recording medium, have long been employed to collect intelligence required to identify and locate targets in support of this tactical interdiction mission.

Limitations of these sensors in support of fast moving tactical combat operations have been multi-faceted. The primary limitation of optical sensors is daytime, clear weather operation. In the late fifties imaging infrared and radar sensors were developed. Problems relating daylight and all weather tactical collection capabilities were alleviated. However, film based collection continued to severely limit the timeliness of the tactical intelligence information collected. For example, the time from recce aircraft over target to target intelligence included aircraft return to base, recce film downloading, film processing, and target identification and location. The time required to accomplish these functions is measured in hours. Increasing mobility of tactical weapon systems dictated that methods be developed to detect, identify and locate tactical ground targets minutes, and even seconds from collection. To accommodate this urgent requirement, near real time image intelligence systems were conceived and research and development objectives identified. The term near real time (NRT) intelligence was coined and defined as the process of providing tactical intelligence to command and control elements within minutes of recce aircraft time over target. To achieve NRT intelligence, revolutionary recce systems were conceived. However, initial concepts (1960's) were technology limited. NRT systems of the 60's employed advanced all weather day/night sensors systems providing an image reconnaissance collection capability which was not limited to fair weather or daylight. However, NRT exploitation of this data was not possible. Extremely high data rates from imaging sensors, limited the recording media to silver halide film. Attendant chemical processing and film handling problems extended the image exploitation cycle to hours.

In the Vietnam War era reconnaissance systems were developed which employed various types of exotic film processing strategies which were designed to speed up the image exploitation cycle. These systems met with limited success primarily because of the specialized procedures employed to process the film. However, time constraints related to analog film exploitation continued. It was apparent that technology was required to replace analog film as the recording media to meet NRT time imagery intelligence requirements.

The first NRT image reconnaissance was developed and experimentally tested by RADC in the late 60's. The system employed an infrared sensor which was simultaneously recorded electrically on tape (eliminating the analog film record) and simultaneously displayed in the aircraft on a CRT display. Real time constraints coupled with limited resolution limited the value of the aircraft display. The tape recorder was, however, successfully exploited at the ground station. With the improved time lines of this system, numerous versions of systems designed to improve timeliness of tactical intelligence were developed. Initial success with electrically recording infrared imagery lead to the development of data link systems to transmit analog electrical infrared and radar imagery directly to the ground exploitation system. This eliminated the time required for the collection aircraft to return to base and download the sensor record. The performance of these NRT sensor/data link systems continues to be progressively improved.

There are three essential elements to a NRT tactical imagery intelligence system. They are NRT time imagery collection, NRT air to ground imagery transmission and NRT imagery exploitation. Systems and technology to field the first two elements are progressing. The emergence of the NRT collection/transmission technologies have magnified problems in NRT imagery exploitation. Time required for target detection, identification and location (NRT imagery exploitation) represents a serious deficiency and limit on the overall effectiveness of NRT image intelligence systems.

NEAR REAL TIME EXPLOITATION TECHNIQUES

NRT imagery exploitation (for the purpose of this paper) is defined as the process of detecting, identifying and locating ground targets within five minutes of receipt of digital imagery from the collector.

The requirement for more rapid imagery exploitation response has resulted in a major thrust of research and development programs in the direction of aiding the image interpreter. NRT image exploitation presents two major inter-related problems to the system designer. They are the high system data rate required to process digital imagery and the time allowed to accomplish the digital image exploitation. The time element is primary. The system must be designed to optimize the time element. When and if a collection platform is directed to a single target or a limited number of targets, data rate is relatively low and the design of a NRT exploitation system is relatively straight forward. The real world situation, however, has shown that NRT collection systems collect large volumes of ground coverage necessitating extremely high data rates. For example, a two minute forward locking infrared (FLIR) mission covering four target sites may include up to 3600 frames of data. Imaging radar systems can collect over 4000 square miles of ground coverage an hour while current manual exploitation averages 200 square miles/hour for the average interpreter using a normal light table and range to higher values for highly skilled interpreters using specialized interpretation equipment. To further compound the problem, more than one collection aircraft may be data linking to a single exploitation station. It is, therefore, unreasonable to even consider brute force (image by image) processing of NRT image data. Traditional approaches which use manual exploitation techniques cannot accommodate the NRT imagery data rates. Methods must be applied to filter and sort NRT imagery so that the target area is detected and extraneous data disreguarded.

Techniques for filtering imagery may be generally sorted into two major categories. The two categories are:

.   Imagery/Intelligence Correlation

.   Automatic Image Processing

Imagery/Intelligence Correlation subsystems involve the correlation of intelligence from other sources and sensors to identify areas of interest on NRT imagery. In most situations the indications from other sources are adequate to detect gross areas of activity. However, the nature of the other sensors is such that target identification and precise location data must be extracted from image source data.

The Advanced Sensor Exploitation (ASE) program is a USAF advanced development program directed at the requirement for technology development in support of imagery intelligence correlation. The ASE program addresses the requirement for the rapid correlation, integration and display of (near) real-time combat sensor information pertaining to high priority mobile ground targets.

The overall program objectives are to develop, demonstrate and validate automated processing techniques/technology to improve the timeliness, completeness and accuracy of dynamic ground target information for tactical command and control elements. The developments focus on the use of advanced (near) real-time standoff sensor/strike systems and the technology required to effectively and optimally exploit these resources in a collective manner in support of the current operations function of the tactical Air Force.

Due to the evolutionary and rapid developments in sensor technology and concepts in this area of (near) real-time combat information processing the program consists of two major phases. The first phase is concentrating on the technology of automated data handling, correlation, tracking and display of the dynamic ground target situation.

The initial demonstration will utilize simulated sensor inputs collected against a simulated enemy force. The successful demonstration of on-line exploitation of the simulated sensors will be the basis for the planned second phase of the program which will transition the technology into a field model for demonstration with live sensors during FY84/85 time frame.

ASE processes and correlates NRT sensor data from advanced outyear airborne sensor systems that are currently under development. Live sensor data from these advanced sensors is not readily available at this time. The ASE approach is to simulate these sensors. Sensor systems to be simulated are from four generic categories: wideband emitter detector; moving target indicator; narrowband emitter detector; and an advanced Synthetic Aperture Radar Sensor. A battlefield scenario is generated by feeding laboratory generated scenario information to the sensor models. Activities focus on developing the high speed automated data handling and display capabilities necessary to fully exploit advanced sensor systems in a complementary fashion; the development of correlation software to present a composite picture of second echelon mobile ground targets from multi-sensor data inputs; demonstrating continuous tracking of these mobile ground targets; developing automated sensor cueing techniques to allow handling off of a group of targets from one sensor system to another when tracking has been or will be lost by a particular sensor; evaluating the digital cartographic data or electronic map technology, required to support advanced sensor data exploitation; and utlimately, developing the capability to maintain and update a near real-time dynamic ground order of battle. The resulting output of dynamic ground target information from the technology developed in ASE is expected to provide C^3I nodes with a composite picture of the second echelon ground environment at a level of detail that will favorably impact target nominations and the interdiction process.

In summary, the objective of the ASE demonstration effort is to develop the automated data processing technology to:

(1) Automatically correlate four generic high capacity sensors: MTI, narrow and wide band emitter detectors and an NRT imaging source.

(2) Demonstrate bi-directional interface with the sensors and tactical command, control and intelligence elements.

(3) Demonstrate automated cueing and tracking of high priority mobile targets over a 200km x 200km geographic area.

(4) Demonstrate display of dynamic ground target situation for tactical battle field surveillance and support to current operation activities.

(5) Demonstrate application of automated imagery/intelligence correlation techniques to filter NRT imagery.

This demonstration will be conducted utilizing a simulated divisional ground force operating over a 200 km x 200 km geographic area. A set of advanced sensor models will provide high volume sensor information to demonstrate the above ASE capabilities. The results of this effort shall be a quantified processing baseline for exploiting and integrating the high volume information available from these advanced sensors.

This advanced development effort is an evolutionary step in the development of automated data handling, correlation and display of (near) real-time mulitsource combat information.

The ASE systems multisource source correlation capability will provide an efficient method of filtering NRT imagery at the location and or situation where other sensor source data is available and practical. In remote areas where NRT image systems are deployed in an austere environment and where complex computer processing capabilities as are typified by an ASE capability are not available; alternate means for filtering NRT imagery must be employed. Since the use of a large number of interpreters to screen and detect targets is not an acceptable alternative; the only alternative is to employ automatic image processing techniques to filter imagery for NRT target detection.

Automatic image processing for imagery filtering can be subdivided into two categories, automatic image screening and automatic change detection. Automatic image screening employs automated techniques for logically processing imagery to determine the presence of target type patterns in imagery. Early attempts at automatic screening and change detection employed complex optical filtering techniques for detecting and identifying target patterns and changes. Size, complexity, and related problems resulted in very little success with this approach. False alarm and missed target rates were unacceptable. The technology explosion in digital processing hardware lead to the application of digital techniques for digital image processing. Digital techniques are currently employed for automatic image screening and for automatic change detection. Description of each of these image filtering techniques follows.

The approach for automatic image screening for NRT target detection is based on technology that has evolved from guidance systems used in air to air rockets, infrared heat seekers, and recent developments in digital target detection techniques that utilize pipeline processing technology to accomplish the target detection. This approach is cost effective using firmware and not software. In the case of infrared the thresholding of signals allows the detection of warm targets against a cooler background and

thus target detection becomes possible, however, for electro-optical and radar sensors the problem is more difficult. The baseline for target detection in real-time or near real time requires use of pipeline technology that automatically performs the same functions on all of the image in a pipeline fashion. This requires that the data be normalized first by a look-up table and at throughput rates such that the detector can be operating on data optimized for target detection. The next process in the pipeline is usually some kind of filtering either low pass or high pass filtering to perform a screening function by eliminating unwanted data bits and extraneous signals. The actual target detection may consist of a gradient analysis and polygon fit with select-able thresholds. The thresholds can be in intensity, area and shape. The setting of automatic threshold control circuits determines the false alarm rate and/or the percent of missed targets. This approach is supported by technology efforts in industry and universities that provide for development of VLSI chips that can be applied to the target identification task in the future.

Target detection is accomplished on the whole image format by using real time decimation and nearest neighbor interpolation processes to reduce the image to an optimized scale for detection but without losing the resolution required for target detection. The decimation process should be selectable with a range from 1 to X pixels.

The target detection would take place directly from the main data stream of digital imagery data prior to storing the full resolution data in disk. This will result in the target detector providing a cue to the control computer that a target exists at a given line and pixel value and that a sub-image containing that target should immediately be passed to the first available interpreter for verification. This eliminates the need for interpreters to screen the complete image for the detected target.

It is planned that automatic identification of targets will be accomplished by carrying the pipeline process several steps further by employing a multi-dimensional filtering capable of performing the identification task. The outputs of the identifier would be annotated at time of verification.

In the case of poor target detection results the interpreter can select a screening mode that will automatically change sub-images for him and allow him to screen and detect targets with computer assistance.

An alternate concept for filtering NRT imagery employs automatic change detection. RADC had been involved in the development of change detection techniques for Synthetic Aperture Side-Locking Airborne Radar (SLAR) since 1965. Computer simulations, using large general-purpose computers, provided the feasibility of automatic digital change detection between pairs of SLAR film images in the late 60's. These studies lead to developed digital procedures for registration of the film pairs that were more exact than any analogue capability. The simulation studies led to the development of a hardwired processor to perform the matching and change detection algorithms. This system, known as the ARRES system, was delivered to RADC in August of 1972, and de-monstrated, on-line in October of that year.

As a result of this success various refinements and additions to the simulation software indicated a need for a change detection system with the flexibility to adapt to algorithm refinements. To meet the new sensor developments a breadboard special-purpose processor was built and demonstrated in January 1973. An advance development model, known as the Digital Modular Change Detection (DMCD) system was built and delivered in 1977. This system is currently being modified to provide a demonstration of automated advance large area exploitation techniques for the UPD-4 high resolution SLAR sensor in USAFE in 1981.

These automated systems for filtering NRT imagery have been touted by the R&D community as the process for processing high volumes of imagery to screen and detect target activity. Successes to date have been limited to laboratory demonstrations under controlled conditions. Results to date indicate problems in two areas:

- False Alarms

- Missed Targets

False alarms are situations when the target detector or change detector indicates targets where no target exists. In an attempt to reduce false alarms, research activity has been directed at techniques for utilizing logic and automated pattern analysis techniques to verify target areas and reject false alarms. Another approach to reduce false alarms has been to increase thresholds and criteria for target detection. This approach, however, also increases the probability of missed targets. (Targets not detected in the autodetection or change detection process.) Although operational unproven, these technologies held much promise for future NRT imagery exploitation systems. With further development image screening devices will be adapted for in-flight application. In airborne application the filter would be applied in the airborne platform, significantly reducing the air to ground imagery data link data rates. However; before airborne or, for that matter, ground applications are considered; very careful analysis of operational test data for false alarm rates, missed targets and other related issues must be accomplished.

To meet these objectives, the Data Handling Recording System (DHRS) is currently under development at USAF/RADC. With the aides designed within the DHRS the equipment is uniquely adapted to performing NRT interpretation of the imagery data. The basic DHRS reporting facility is essentially composed of four distinct parts or modules (refer to figure 1): SIM (Sensor Input Module), S/RM (Storage/Retrieval Module), RTPM (Real Time Processing Module), and NRTEM (Near Real-Time Exploitation Module). The digital imagery data will be stored in the S/RM with the digital image represented as a complete entity.

A ground target screening system will provide for the detection of man-made objects and for the classification of targets based on geometric characteristics (e.g., edge data, object brightness, symmetry). Classification of objects (e.g., tanks, trucks) is based upon object size, texture, contrast, edge straightness, and components (e.g., wheels, tracks). The extent of target classification that can be incorporated into the Real Time Processing Module will be limited by the state-of-the-art of that technology. The classifier portion of the Real Time Processing Module will be designed so that modification (software and hardware) can be incorporated as technology advances. RTPM will (upon the detection of a man-made object) extract the portion of the frame in which the object is located for further analysis (classification and display). The object in the displayed scene will be highlighted so that it can be detected by the display operator for target confirmation at the exploitation module. The highlighting symbol (alpha / numeric) appearing on the display will indicate the type of target detected. The Storage and Retrieval Module will provide for a full-frame imagery storage interfaced to a full display image for image enhancement. This technique is better than other approaches considered since only pertinent portions of the original (sensor) digital image is extracted and less important portions are deleted. This will give the TAC commanders a greater flexibility in viewing only those portions of imagery that are of importance for timely decision and strike. A computer will monitor the data (I/O) and functions of the processing systems. Such a design scheme will provide a centralized control of the digital storage, enhancement, and dissemination to Tactical Commanders.

The Near Real Time Exploitation Module of the DHRS will provide a high quality image display, alpha/numeric annotation, image enhancement, and dissemination. This will consist of the following equipment:

(1) Digital Image Enhancement Equipment
(2) High Quality Image Display
(3) Image Processor
(4) Alpha/Numeric Software
(5) Dissemination Subsystem

In a typical scenario, the image screening module will designate and determine man-made objects (MMO's) (whether the objects pertain to a specific geometric shape; e.g., circles (oil tanks, drums) squares, and rectangles (buildings). After a desired scene has been extracted from the frame and analyzed via the processing module the image is then transferred to the NRT Exploitation Module. If target confirmation (classification) cannot be determined, the operator will have the capability to pinpoint desired pixels (and subsequent locations) to a particular region of the extracted frame. Alphanumerics/Graphics will provide the capability to annotate the imagery on the imaged display. The pattern recognition software will provide appropriate symbology to identify buildings, roads, target positions, and other distinguishable cultural and terrain features. The Alpha/Numerics (Graphics) will provide the capability of annotating with appropriate symbology particular target areas of interest. The Digital Image Enhancement will provide any required enhancement to the image displayed.

The Display will be the central unit around which a facility configuration will be assembled. This will consist of a high quality CRT display with full and psuedo color capability, zoom and scroll features, and the capability to store selected digital imagery.

The Image Processor provides the display, with the necessary software to enhance the displayed region by correcting for radiometric and geometric distortions, providing filtering functions (high and low frequency discrimination), edge enhancement and aligning provisions. The main CPU (Controller) provides the necessary algorithms for computation of ground coordinates via a priori information; e.g., mapping functions, sensor orientation, and INS (Inertial Navigation System).

Areas such as automatic target screening and visual image enhancement for target detection/classification are applicable to the accomplishment of the target detection function. Emphasis in these areas is on automatic processing of the sensor input data rates. An operator performing the target detection function is not capable of interactively manipulating complex systems to optimize the detection of targets because of the extremely high data rates for raw sensor data. On the other hand, implementation of image processing techniques to assist in performing the target identification function are not subject to this same criteria. The input data rates for the operator would be substantially less and the operator would have time to interact with the system to optimize the visual display of a target.

The primary purpose of the DHRS is to provide design data for a digital system that will provide real-time targeting information and to provide a technical base for the design of an airborne display for use in future recce/strike aircraft. The DHRS will also be used as an evaluation facility for advanced real-time sensor systems and

to test and evaluate false alarm rates, missed target rates, and other issues related to the automatic processing of NRT imagery.

The system proposed will reduce the time now required between intelligence gathering and resultant decision making. The major technical/human interface will be the integration of the components which will reformat and input the real-time data and perform target detecting and image enhancement to aid in the human interpretation of the imagery. The workload of the human interpreter will be reduced to a manageable task of identifying targets which have been detected via the DHRS hardware/software modules.

Future research and development efforts will emphasize the ability to demonstrate actual improvements in target detection and identification. As an example of this approach, RADC (Rome Air Development Center) is currently implementing several programs to test the applicability of image processing techniques to the real time environment. In one program, sequences of video frames are being digitized and processed by real-time algorithms. The sequences of video frames are then evaluated using human interpreters to determine their applicability to the real time environment. Only those image processing techniques yielding concrete improvements in detection or identification will be considered as candidates for future implementation.

The third area of NRT imagery exploitation is target location. A NRT precise target location subsystem is essential to support NRT tactical strike concepts. The current approach to the development of precise target coordinates employs maps and charts or point positioning data bases. Coordinates developed from maps and charts lack precision required to support all weather/night strike systems. The point position ing data base (PPDB) system is currently being used to develop precise coordinates. None of these systems, however, are capable or have the growth potential of being responsive to a near-real-time environment. Storage and retrieval of the PPDB is a manual, time consuming, process. Automatic methods for handling high quantities are not available and would be extremely expensive to develop and maintain in operational use. Point transfer is a manual or computer-assisted process and currently is both time consuming and prone to error. The time required for a point transfer alone in some areas can consume 5-15 minutes for imagery selection and system set up and 5-10 minutes for location a single point. To achieve the NRT goals, the location function must be accomplished in seconds and minutes and not impact or delay the detection, identification or reporting processes.

The technical approach for implementing an advanced NRT precise target location subsystem is directed at achieving the program objective of deriving, in near-real-time, the precise location information of potential targets with enough accuracy to support (1) identification, (2) the decision to strike or attack, and (3) the execution function, including the high accuracy required for night and adverse weather tactical weapon delivery systems. The approach allows for the development of precise locations of any or all targets from multisensor imagery.

The basic location subsystem will be configured around state-of-the-art technology, however, a modular approach to implementing the target location subsystem will permit the flexibility to incorporate new techniques and technology as they are developed.

The advanced target location subsystem development provides the capability to derive precise target location information in near-real-time. An all digital approach has been selected that will provide the location information on-line or integrated with the functions of detection and identification with the coordinates automatically incorporated into the report. The photo interpreter will denote a detected target on the display subsystem with a cursor or light pen and command the system to determine precise location. The target location subsystem would then accomplish the necessary actions to derive the locations in whatever coordinate reference system is desired, and provide the coordinate information back to the operator for inclusion into the report.

Target location will be accomplished by utilizing proven photogrammetric techniques and some yet-to-be developed technology. Two major methods are feasible depending on the inputs and the accuracy of those inputs available from reconnaissance imaging sensors. If the reconnaissance sensor data include very accurate vehicle position and attitude information and the math model of the sensor is known to a high degree of accuracy; a direct target location technique can be developed to generate accurate target coordinates very rapdily. However, if the sensor vehicle position, additude, and math model information is not available or is not highly precise; then a technique that relies on a point positioning data base (PPDB) must be utilized.

Both techniques (PPDB dependent and PPDB independent) are being developed because it is highly probable that the precise vehicle position and attitude information for all of the sensors will not be available for some time. Technical descriptions for each of these techniques follow:

The point positioning data base dependent technique is comprised of three subsystems. They are, a digital point positioning data base, a PPDB storage and retrieval device, and a digital target location module.

Exploratory development efforts underway at RADC are currently addressing the specific required characteristics and will develop an experimental PPDB for subsequent use in NRT target location demonstrations. The point positioning data base will consist of digital image data, parametric data, digital terrain elevation data (MATRIX type elevation data) and possibly digital cartographic data.

The PPDB storage and retrieval device is a subsystem for storing of massive amounts of digital image data. A typical point positioning data base will be expected to cover approximately 100,000 square miles which necessitates a very large storage capability. In order to meet the throughput response objectives of the program, this digital image data must be randomly accessible in a matter of a few seconds. The data transfer rates must also be very high. A device will be designed, fabricated and interfaced to the target location subsystem processor that will provide the necessary storage and retrieval of the PPDB. Optical disk technology appears very promising for this application.

The digital target location module will be composed of several sub-modules that provide digital image viewing/mensuration, a point transfer capability, processor(s) and software. The viewer/mensuration operator's station device may be the display subsystem used to support NRT target detection/identification or may (of necessity) be another viewing device. The viewer must provide a display of reconnaissance imagery to the operator for determination of the pixel location of the target within the overall image display. The pixel location information would be processed using the sensor geometry position and attitude information to determine approximate ground coordinates. These coordinates would in turn be used to retrieve the appropriate PPDB image from the storage and retrieval device. The PPDB would be displayed and the operator would then perform a manual or a computer-assisted point transfer. The ultimate objective is to develop an automatic point transfer capability. This capability would permit point transfer of the target on the recce image into the PPDB to be accomplished much faster and would be completely automatic. With the point transfer accomplished, the processor would then perform the photogrammetric calculations to derive the ground coordinates. The final derived coordinates (x, y, z) are then provided to the operator and provided to the report generation subsystem.

The PPDB independent target location subsystem, also known as direct target location, is addressing the capability to derive target location information without the use of a PPDB. Direct target location can be accomplished when the attitude, position, sensor geometry, and image plane coordinates are known accurately enough to define a ray in space and intersect it with the earth's surface. The PPDB Independent capability would be primarily a software module that could be incorporated into the system. Only the designation of the target on the display screen to the operator would be necessary to generate accurate coordinates. Direct target location is complementary to PPDB targeting and covers the contingency when PPDB's are not available.

SUMMARY AND CONCLUSIONS:

The advent of digital technology to support NRT image exploitation offers the opportunity to develop and field a "universal" digital image exploitation system. The "universal" system would employ a "standard" sensor interface such that all NRT digital image sensor systems would be supported on the ground by a common image exploitation system. This common system could operate relatively independent of sensor type, source or even nationality. Software modules would be used to configure the common system for the sensor to be exploited. Future image sensors could also be accommodated in this way.

The common image exploitation system offers other advantages. As automated digital image pattern recognition technology is developed and tested; the common system could be modified in a modular fashion to take advantage of these advanced techniques.

Finally, and of primary significance to NATO, the common system concept is ideally suited to supporting NATO forces in that the ground exploitation station can be made relatively independent of sensor collector type or nation of origin. The flexibility of the common system concept also provides opportunities for NATO cooperative/joint research and development programs dealing with high technology needs to support NRT image exploitation. The following are potential technical areas where this cooperative/joint R&D could be mutually lucrative.

. High Resolution Displays

. Automatic Target Detection

. Automatic Target Identification

. Multi-Sensor Correlation

. Automatic Change Detection Techniques

. Standard NATO Sensor/Data Link/Ground Station Interface Definition

. Standard/Compatibility NATO NRT Report Format Definition

This list is not intended to be all inclusive. It is suggested as a catalyst to promote advocacy of joint NATO program which would ultimately result in a NATO NRT imagery exploitation system. Comments relating to mechanisms for initiating such a program and pros and cons for a common NATO image exploitation program are encouraged.

A SYSTEM LOOK AT REAL TIME PROCESSING FOR TARGET ACQUISITION AND CLASSIFICATION

Henry Lapp
Air Force Wright Aeronautical Laboratory
Avionics Laboratory
Wright-Patterson AFB, Ohio
U.S.A. 45433

SUMMARY

An important role for target acquisition is the location and identification of targets in near real time. Current technology has been compartmented into sensors, processing, air or ground exploitation and finally dissemination. In the days of hot spot or radar blip detection, this segmentation of functions was appropriate. With the current emphasis on real time decision making from outputs of high resolution sensors this thinking has to be re-analyzed. A total systems approach to data management must be employed using the constraints imposed by the atmosphere, survivable flight profiles, and the human workload. This paper will analyze the target acquisition through classification tasks and discuss the machine processing and data screening techniques that are applicable. The data handling capabilities of an on-board operator and ground based image interpreter are compared. A philosophy of processing data to get information as early as possible in the data handling chain is examined in the context of ground exploitation and dissemination needs. Examples of how the various real time sensors (screeners and processors) could fit into this data handling scenario are discussed. Specific DoD programs will be used to illustrate the credibility of this integrated approach.

THE NEED FOR RAPID TARGET ACQUISITION

Since the Blitzkrieg type of warfare was successfully demonstrated in World War II the mobility of ground forces has constantly increased. It is now possible for very large, well equiped, full spectrum, ground forces to overrun defenses and thrust from tens of kilometers to a hundred or more kilometers per day, obtaining and maintaining control of the entire area encompassed, or to rapidly deploy in preparation for an attack. Suffice to say the battlefield, as well as the domain of air operations, can be extremely fluid.

It is the prime purpose of tactical reconnaissance to provide the theater and the field commanders with the necessary information about the enemy's operational situation to successfully conduct operations, both air and ground, in such a fluid environment. Obviously this information must be timely.

When the enemy has the prerogative and the opportunity to conduct offensive operations, the need for timely information becomes increasingly critical. The defense must effectively react while the opportunity is available or before it is too late and their capability to do so is overrun or otherwise negated. In this type of highly mobile warfare tactical aerial reconnaissance plays a vital role, being the most mobile of all capabilities. A reconnaissance aircraft can travel at a rate of 18 kilometers per minute or more, permitting it to reach an area of concern from a remote base in a few seconds. The crux of the problem is to make the information, so quickly gatherable, useful in time for reaction. To date this has been effectively accomplished too infrequently.

Traditionally, airborne reconnaissance has been accomplished by a reconnaissance vehicle flying over the area of interest. A variety of sensors aboard the vehicle sense and record the reconnaissance information for subsequent processing, interpretation and reporting of the digested information. These functions normally occur upon landing at the reconnaissance base. The time involved in this total process, from sensing to reporting, varies from hours to weeks, depending upon the type of missions and the quantity of data sensed (the extent of the coverage and the resolution), and the type of information to be extracted. This time factor can be broken down into functions as follows:

Air Transportation - Time for the vehicle to return to base and land. (Typically 10 min. to 1 hour)

Ground Transportation - Time to transfer the recordings from the vehicle to a ground processing station. (5 to 20 min.)

Data (or Image) Reduction - Time to process the recordings into extractable form. (10 min. to 2 hours)

Target Detection and Location - Time to find probable targets/areas of interest in the total recorded take. (2 min. to hours)

Intelligence Interpretation - Time to interpret the recordings into applicable information. (5 min. to days, or weeks in the case of maps). This usually involves correlation with other information.

Reporting - Time for reporting procedures (Preparation, Dissemination). (A few min. to hours, days for maps)

The above time factors are typical or average values. They may be significantly less for some reconnaissance missions where the number of specific targets to be reconnoitered is small and urgency procedures can be applied throughout the cycle. For missions in which the information perishability factor is an hour or more, traditional capabilities are adequate to solve the problem. They will continue to provide the bulk of the reconnaissance and intelligence information obtained. However, for reconnaissance of more perishable, or more urgent targets the normal reconnaissance-intelligence cycle must be drastically compressed or significantly changed into a real or near real-time reconnaissance capability in order to be responsive to a highly mobile threat.

The functions of target detection, location and interpretation normally involve the human eye and brain processes which are not amenable to significant time compression. Consequently, for near real-time operation, those reconnaissance jobs requiring sub-stantial human study must not be undertaken or the process will rapidly outgrow its required time frame. The amount of information for which the human is tasked should be no more extensive than that absolutely required for the primary purpose of the mission. Only answers to simple and precise reconnaissance questions can be expected from a human observer in real or near real-time. It follows that real and near real-time methods should only be applied to obtain the minimum information essential to make a decision where rapid response is likely to be required.

The historical method of gathering data is with a film camera. A film/camera data rate of 10^9 is not uncommon. For the real-time case a sensor data rate of 10^6 to 10^7 pixels per second as experienced in typical tactical scenarios. This pixel rate number is arrived at from convoluting the ground resolution, swath width and V/H necessary to adequately perform the tactical mission. The human in the interpretation task has a 10^5 data processing rate. (Gardiner and Nicholson, 1971).

Because of the limited amount of information which a human can handle in real or near real time, and the tremendous amount of data which a sensor can acquire in real time, a human observer in the reconnaissance aircraft may well miss or overlook very important information which may not be quite as time sensitive as that of initial prime concern. For this reason, as well as for verifying the airborne interpretations, it is important that permanent recordings be made of all sensed information for more thorough assessment after return to base, even though the intent of the mission is one of obtaining a specific real or near real time reconnaissance answer.

OPERATION ASPECTS OF REAL-TIME RECONNAISSANCE

The traditional reconnaissance information process cycle was broken into steps and defined. They are repeated here for convenience:

1.	Air Transport	4.	Target Detection and Location
2.	Ground Transport	5.	Intelligence Interpretation
3.	Data Reduction	6.	Reporting

The use of a wide band data link makes possible the elimination of several of the above reconnaissance cycle steps and the possible reduction of the time factors of others. For example, if the data link can operate directly from the reconnaissance vehicle, while it is acquiring the information, and can transmit at the acquisition rate, then steps 1 and 2 above are eliminated. In addition, step 3 can be reduced by operating the recorder-processor in tandem with the receiver, at the information acquisition rate. Rapid film processing technology can reduce step 3 to under 2 minutes, under special conditions to as low as 10-15 seconds. Even if the reconnaissance aircraft cannot transmit while acquiring, a buffer storage can be provided in the aircraft to permit transmission when the aircraft reaches a position from which it can transmit. This will not eliminate step 1 but will reduce it considerably; step 2 will still be eliminated; and step 3 will be reduced as above. Step 3 can be reduced to zero by use of a dynamic display of the information instead of, or in parallel to, the recorder.

For steps 4, 5, and 6, the human is normally involved and reduction of time elements for these steps can only be reduced by changing his duties. The time required is a function of the area covered by the reconnaissance per unit time, the number, size and deployment of targets of concern, other target characteristics such as smoke, dust, or tracks, the ability to extract the pertinent information from the scene (clutter, contrast), the number of interpreters employed, the skill and experience of the interpreters, and the procedures used. The target/background characteristics (such as clutter and contrast) are always dominant factors.

Real and near real-time transmission of reconnaissance imagery can thus reduce the total time factor between acquisition of the reconnaissance and reporting to a matter of approximately 10 minutes up to an hour or so. Wide band data transmission is thus most useful for information whose perishability factor is in this time range. However, this conclusion is valid only when the human's time factors do not dominate the equation. Otherwise, the process becomes swamped, quickly breaks down, and is useless.

For operation in a high jamming environment or line of sight conditions where a wide band data link capability may be unuseable, the same results can be achieved with a dynamic display for the reconnaissance crew in the reconnaissance vehicle and having them perform the entire process and report via a narrow band link. The reconnaissance cycle time is thus reduced to the time required for the reconnaissance crew to detect and recognize (interpret) the targets and verbally or symbolically report their findings. Within certain constraints this can be a matter of seconds.

If the reconnaissance crew is delegated the authority to make strike decisions, it is now a reconnaissance/strike capability. This reconnaissance/strike process has been accomplished for many years for a limited range of applications and under restrictive conditions. The traditional forward air controller, operating in low-performance aircraft, is an example of the entire process which usually includes decision making and direct or indirect control (target designation) of strikes. Until recently the principal reconnaissance sensor of the forward air controller has been the human eye, unaided or aided with various types of visual optical sights. In addition, a substantial amount of reconnaissance information has been collected visually by strike pilots as an adjunct, or in addition, to their primary strike missions. In these cases the process has been in real or near real time although no reconnaissance sensors nor displays of pictorial reconnaissance information, per se, were necessary. However, the conditions of operation have been restricted to daylight and good visibility.

Sensor technology now permits the target acquisition capability to be extended to nighttime operation and/or conditions of visibility well beyond the eye's unaided capability. In addition, depending on the scenario, the near real time target acquisition capabilities may be accomplished from high performance aircraft. In these cases the target rate, clutter and contrast become more and more important. While much work is progressing to model and quantify the effect of the physical variables on the human performance of detecting and recognizing targets at real time rates, the variables are so complex, numerous and interrelated that the problem cannot be generalized. The remainder of this paper will discuss techniques to assist in the real/near real time performance of sensors/humans to perform target acquisition missions and the operational limitations of such techniques.

One of the prime functions of aerial reconnaissance is to maintain periodic surveillance of a large area of concern to alert the commander to enemy activity or movements. As these missions must, of necessity, be flown at high altitudes in order to survey wide areas, line of sight to the receiving station (ground or airborne) can be achieved. Use of directive and tracking antennae reduces the susceptibility to jammers not directly or nearly in the line of sight. Only an airborne jammer nearly in the line of sight can effectively jam the receiver and thus the transmission.

For the surveillance mission, where activity is the prime factor of interest, targets need not be recognized as such. Just changes or movements or other unique features of detected objects need be observed. Of course, the objects detected must have a high probability of being of military significance, or this relationship must somehow be deducible. For this purpose cues can be used. The most readily available cues are:

1. Motion (MTI Radar)

2. Change detection (Radar)

3. Electromagnetic Emissions

4. Heat Emissions (Infrared Sensor, weather permitting)

Obviously the degree of normal, non-military, activity of the surveyed area must be relatively low. A convoy of military vehicles on a busy highway will be detected but only supplemental information will differentiate them from normal civilian traffic. On a little used secondary road, however, the appearance of many moving objects would be cause for alert.

In any case the decision of where and how to react will probably require a specific reconnaissance mission to determine, through recognition, what the motion or activity really means. The most useful purpose, then, of a surveillance mission is to highlight points or small areas for more detailed coverage, as well as to provide alert warning. As this mission is accomplished from high altitude and covers large areas, radar, either MTI or side looking-snythetic aperture, or more desirably a combination of both, can provide an all weather surveillance capability of wide areas for activity indication and alert purposes. The information, either raw or partially processed, may be linked to the ground via wide band data link and automatically processed into imagery for manual read out. If provided in digital form, the information may be automatically compared to information from a previous pass of the same area in a digital change detector with all changes superimposed on the radar map. Such a radar surveillance capability may be augmented with electronic reconnaissance and/or infrared for the detection of targets with those specific characteristics. For an analog system the human is required to look at imagery and, depending on the situation, could soon take so much time as to obviate the advantages of real-time transmission. Automatic digital change detection can overcome this limitation, and provide the necessary movement information within the allowable time.

For detail examination of specific points, small areas or routes which have been highlighted by surveillance missions, or other sources, as requiring closer scrutiny, specific reconnaissance sorties are needed to provide the necessary information. Today's state-of-the-art in automatic target identification is not sufficiently advanced to preclude the need for human recognition of targets on which to base strike decisions or tactical employment commitment. This is not to say that exigencies of the situation may not require action based on such a dearth of confirmatory information. The operational situation may be critical enought to "blast every blip in sight" but in the absence of the need for such desperate measures it seems fair to assume that commanders will require more positive identification, based upon recognition or at least reliable classification, before major commitment of resources. Automatic classification techniques, discussed later, are under development, but sensors providing sufficient information to the processor, plus processor accuracy, completeness and low enough false alarm rates for operational reliance have yet to be flight demonstrated and proven.

As humans must, as yet, do the recognizing of targets this must predominantly be done by virtue of the shape and size of the target. Today's radars or microwave sensors do not provide sufficient resolution for recognition of most tactical targets by their shape. Their resolution is improving (with complications) but many unanswered problems, such as specularity, susceptibility to electronic countermeasures, etc., need resolution. Thus, the primary recognition sensors (other than direct eyeball) must now be electro-optical or infrared imagers which present visual-like information to the human. Operating in the optical wavelengths (less than 300 micrometers), these sensors are susceptible to atmospheric degradations. They cannot operate through clouds so, in the presence of cloud cover, they must be used below the ceiling. Also, due to atmospheric haze, sensor performance improves as the distance to the target is decreased. Sometimes (in heavy fog) they are totally ineffective.

The atmosphere factors, plus the current trend of flying very low to survive in today's high threat environment, dictate that this type of reconnaissance mission be performed at low altitude.

Low altitude operation over hostile territory poses two very severe problems for wide band data links. Such links require line of sight from the transmitting aircraft to the receiver. If the transmitting aircraft is at low altitude, terrain masking prohibits the maintenance of line of sight to a ground receiver. The receiver, either as a relay or as the interpretation station, must be at high altitude. Such a receiver is highly susceptible to jamming by very unsophisticated means. Its overall operating cost is high and it is a lucrative target for direct attack. Even if the receiving station employs highly directional antennae tracking the transmitting reconnaissance aircraft, if the jammer is anywhere in the vicinity of the reconnoitered target area, then the jamming can only be overcome by one of the following:

a. Recording the reconnaissance data for transmission when the reconnaissance aircraft proceeds away from the jammers (if it climbs to high altitude then it can transmit directly to the ground, as does the surveillance mission, and a relay if not required but the highly directional tracking receiving antenna is).

b. Transmitting a moderate information band width, over a very high transmission band width digital data link while using sophisticated digital processing techniques to overcome the jamming. This restricts the transmitted information band width to, at best, several hundred kilobits per second (a high altitude receiver with directional tracking antennae is still required).

c. Observing the target areas and interpretation of the target situation by the reconnaissance crew with direct or subsequent reporting over a narrow band width, jam resistant (possibly secure), digital data link by voice, symbology, or very limited pictorial coverage.

REAL TIME PROCESSING FOR TARGET ACQUISITION AND CLASSIFICATION

Cueing for the wide area surveillance mission has already been discussed. Use of autonomous cueing, screening or target classification techniques on specific reconnaissance missions would enhance the search capability considerably to accomplish reacquisition of targets which have moved, or to overcome positional errors in either the original target location or in the navigation. They may even provide some capability for finding targets of opportunity. If the cueing technique is automatic the reconnaissance crew or the ground interpreters are assisted in the search function (a function which the human does very poorly but machines can do well) and permits their concentration on the recognition and identification functions (which machines do poorly but humans do well). Several cueing techniques will be discussed below, along with their applications in conjunction with an identification sensor to provide a localized search/identification capability. These cueing techniques are highly sensor, atmosphere and target dependent. Specific unique target characteristics are used and false alarm (non-targets with similar unique characteristics) rates are constantly a problem. While clutter, without target-like characteristics, is no problem, in a highly target rich environment cuers can become saturated. In this case, selection of a particular cued target for identification must be either manual or random.

CUEING USING THE IDENTIFICATION SENSOR

Infrared cues. Under many conditions certain types of targets emit strong infrared radiation by virtue of their temperature/emissivity characteristics. These emissions may be detected at long range by a FLIR operating in the "wide" angle mode near the horizon, although target masking by terrain or foliage may be a problem. The sensor may be "panned" slowly across the forward horizon for the detection of such emissions. Upon detection the narrow angle of the FLIR may be employed, centered on the "hot spot" and closure employed until recognition. In the case of a downward looking infrared sensor, hot spots on the imagery automatically draw attention to specific points and their immediate surrounds. While hot spot cueing is inherent in an infrared sensor it is not effective if a prolification of hot spots appears.

Automatic Screening. Considerable work, with some success, has been accomplished on applying automatic screening techniques to analog video signals, especially from infrared sensors (both line scan and FLIR). Simple algorithms have been developed to detect targets, with the first iteration being on edges and lines characteristic of man-made objects and further refinement being made on size and geometry. The screener indicates which areas do or do not contain man-made objects. The further refinements discard such man-made objects as roads and large buildings. Detection probabilities on vehicles have been high, but false alarm rates are highly dependent on the background due to the fact that features of vehicles (size, aspect ratio, etc.) used in the screening process are the same as those of some non-vehicle man-made objects of no interest. The principal virtue of this technique is that, in real time, it automatically discards areas of a scene wherein the probability of there being a man-made object is low, thus improving search time by a human. However, in a highly cluttered urban scene very little area is discarded, limiting the usefulness of this technique to rural or natural backgrounds. The screener output may be superimposed symbolically on the observer's display and/or record either in the cockpit from a FLIR or line scan sensor, or at the receiving station, to highlight those small areas where the probability of there being a target is high.

The Avionics Laboratory is currently pursuing a more advanced form of screening technology. The early thermal sensors provided the observer sufficient resolution to detect hot spots. As thermal sensor performance improved, the image quality became adequate to operate on the external features of the targets (edges, length/width, shape). The screener technology was considered an add-on to a sensor designed for a human observer. The current Avionics Laboratory thrust is to let the screener/processor requirement drive the sensor design. Sensor trade-offs will be made in the areas of field of view, dynamic range, resolution, sensitivity and image enhancement.

The number of detectors on the focal plane has been limited to some degree by the practical limits of the display; 525 or 875 lines on the display were the two most popular sizes. The limit was really the human observer's inability to cope with the band width fed to him from the display. The other driving limit on number of detectors was the desire to maximize the field of view to allow the observer to pilot the aircraft. Now, with the introduction of image screening and the advances in butted infrared focal plane arrays, concepts for sensors with thousands of detectors are conceivable for the piloting task. The large focal plane can be used to display, at much reduced resolution, a wide field of view. For the target acquisition task, the large focal plane allows sufficient resolution to be achieved over a wide field of view. The screener, operating on this full resolution, can search the larger area and cue potential targets. The cued targets can then be highlighted on the low resolution wide field of view display on a blow up of the target and be displayed to the observer at full resolution.

The image screener is able to handle real time band width orders of magnitudes greater than those handled by the observer. It is now practical to design the thermal sensor to give an output that allows the screener access to more target information than would be presented to the observer. The dynamic range of Electro-Optical multiplexed FLIRs is in the 20-30 db range. Digital multiplexing would allow the dynamic range to expand to 50-70 db. The additional information cannot be displayed but can be utilized by the screener. The same logic is true for types of enhancements. The human and the processors have different needs. The conclusion of this trend is that it is now possible to design a sensor/processor system that provides imagery of sufficient quantity that the screener is able to do a much improved job of false alarm rejection or target classification. Syntactic image screening is one example.

Syntactic screening means using the geometric relationship of the internal target structure to provide more information for decisions. A tank has a hot engine compartment, hot exhaust ports and sometimes, hot treads. The geometric relationship of these heat sources is known. If the heat sources are present, the screener can utilize them to further improve the probabilities of correct classification. In the past, we taught the image screeners to recognize external shapes such as circles or triangles. Now we can use the improved sensor thermal or special resolution to operate on the internal shapes. The screener can be taught a circle on top of a triangle is an ice cream cone. In cases where part of the target is masked, this internal structure is useful for separating potential from false targets.

CUENG USING A SEPARATE SENSOR

Radar. If the specific reconnaissance aircraft has a radar capable of detecting military targets plus an MTI capability it can provide cues on which to point the TV or FLIR sensor for recognition and identification or to mark the line scan imagery. The radar must be capable of acquiring targets at low depression angles but at moderate ranges (less than 10 Km) from low altitudes. It must be capable of detecting the same types of targets that the surveillance radar does in order to reacquire those which have moved in the interim. At low altitude operation at very low depression angles foliage masking becomes a severe problem in many geographical areas. This problem may be overcome by using long wavelengths (low frequencies) or combinations of two or more long wavelengths. Such techniques have been proven to detect military targets in heavy foliage. This radar need not look forward but can operate in the side looking, synthetic aperture, digitally processed mode. As it need only detect, not recognize, its resolution can be poor and it need not "map" the terrain, just provide the direction and range to point the identification sensor or mark the line scan imagery. At low resolution, band width is low and cost can therefore be quite low compared to most radars. The TV or FLIR can then be zoomed in on the suspected target by virtue of approach and by narrowing the field of view with the operator concentrating on identifying, not searching. A finite number of false alarms can be tolerated as long as principal targets of interest are detected and the clutter is not too great.

The Avionics Laboratory is currently evaluating a digitally processed low frequency SAR radar, called IMFRAD (Integrated Multi-Frequency Radar). Preliminary results indicate the ability to detect tactical targets masked by camouflage and foliage from low altitude standoff surveillance ranges. The IMFRAD radar information is processed on board the aircraft in real time and does detect both stationary and moving tactical sized targets. Because of the automatic processing (the radar detects only the metal vehicular sized objects) the IMFRAD type radar can be used to search large areas and cue a high resolution sensor for the final target classification.

Three Dimensional Target Classification - Recent experiments with a modulated, laser line scan sensor have indicated the possibility of a very discrete automatic target classification capability using phase information to very accurately discriminate specific military objects by virtue of their shape and height characteristics. Automatic, real time classification of specific types of targets has been demonstrated with a high probability of classification and a low false alarm rate. To date, coverage of such a sensor has been restricted to no more than 20° from nadir so its applicability as a cuer for human confirmation with a FLIR or TV in a high speed vehicle, or for very wide angle coverage, remains to be demonstrated. Its output can be directly applied to a downward looking sensor. In fact, being a line scanner, it has its own pictorial output in the reflectance domain as well as a three dimensional image.

The Air Force is embarking on a program titled Target Cueing and Classification Sensor. This program is designed to provide USAF with a real time, airborne, target classification (recognition) capability for the low altitude, high speed, penetration mission. The sensor will be capable of obtaining three dimensional spatial information and classifying targets based on shape discrimination to provide a high probability of detection and a low false alarm rate. The system will process the scene data in real time, <0.1 sec, and provide the user with the target name and location with respect to aircraft position.

This automatic target classification capability will provide usable reconnaissance information in real time for tactical missions. This is especially valuable against mobile tactical vehicles such as SAMs, AAA, tanks, etc. This system will have direct application in the TAC Quick Strike Reconnaissance and Strike Control and Reconnaissance missions as well as in conventional real-time reconnaissance. Since targets are classified and located in real time, information can be transmitted to the user over a low band width data link. The cues will also alleviate the human saturation problem experienced with uncued high band width imagery.

Multispectral Cueing. Use of two or more sensors or one sensor operating in more than one spectral band, or wavelength, can provide information which can help to discriminate targets from their backgrounds. This is particularly appropriate to camouflaged targets in which other, readily sensed, characteristics may be the same. A tremendous amount of research has been applied to spectral signatures and discrimination techniques.

TABLE I

INCREASE FUNCTIONAL UTILITY VIA MULTI-SENSOR

	MMW RADAR	FLIR	CO_2 LASER
FUNCTIONS			
WIDE AREA SEARCH/ACQUISITION	X		
RECOGNITION		X	X
RANGING	X		X
TRACKING	X	X	X
MOVING TARGET INDICATION	X		X
NAVIGATION	X	X	X
TERRAIN FOLLOWING/AVOIDANCE	X		X
WIRE/OBSTACLE AVOIDANCE			X
WEAPON DELIVERY/GUIDANCE	X	X	X
CAPABILITIES			
DAY/NIGHT	X	X	X
ADVERSE WEATHER	X		X
ALL WEATHER	X		
SMOKE PENETRATION	X		X
COVERT		X	QUASI
PROJECTED MUNITION COMPATIBILITY	X	X	X
PROCESSING COMPATIBILITY	WITH CO_2 LASER RADAR		WITH MMW RADAR

Most of the research wavebands in the visible and air infrared electro-optic spectrum. With the increased emphasis on decoy discrimination increased range and improved dust/smoke penetration, the spectrum of interest for multiple wavelength imaging has been extended to the millimeter and radar wavebands. Table I illustrates the utility of the millimeter and active and passive thermal bands in performing various functions under various environmental conditions. The trade-off of "optimum" sensor mix also depends on the state of the development of the sensor technology. The associated processing available and the scenario selected. Table II rank orders sensor suits for the European Theater. Again, the ranking did not consider the state of development of the sensor and sensor integration/cueing technology. Currently, the Avionics Laboratory is pursuing the development of advanced sensors in the millimeter, and active and passive wavebands.

TABLE II

RECOMMENDED SENSOR SUITE PERFORMANCE RANKING FOR EUROPEAN THEATER (U)

SENSOR SUITES	GCI/SAM RANKING	ARMOR RANKING	AIRCRAFT RANKING	OVERALL PERFORMANCE RANKING
ACTIVE/PASSIVE NEAR IR	4	4	2	3
PASSIVE IR FLIR	7	4	2	7
PASSIVE/ACTIVE IR	5	2	3	2
PASSIVE/ACTIVE IR PLUS MMW	6	1	4	1
HIGH RESOLUTION SAR	8	5	1	5
SAR AND EM EMISSIONS	1	6	5	6
SAR AND EM EMISSIONS AND DECOY/DECOY	2	7	5	8
ACOUSTIC AND PASSIVE MICROWAVE	3	3	5	4

In parallel, a Targeting Systems Characterization Facility has been established. This in-house Avionics Laboratory effort operates; measures and models the performance of targeting sensors under poor weather conditions. The ability of various sensors to locate and identify military targets is evaluated over a calibrated atmospheric path. The atmospheric data is also made available to update and improve atmospheric models such as the Air Force Geophysics Laboratory LOWTRAN atmospheric propagation model. Simultaneous EO Sensor/Atmospheric Correlation and Atmospheric Transmission is measured in the visual, IR bands, 10.6 microns, and in the 95 GHz band. The parallel improvement of sensors, processors and actual simultaneous multi-wavelength measurements allow optimization of multispectral sensor suits.

TABLE III

EFFORTS IN CUEING, AUTOMATIC SCREENING AND TARGET CLASSIFICATION

WAVEBAND	ONGOING		PLANNED	
	EXPLORATORY DEVELOPMENT	ADVANCED DEVELOPMENT	EXPLORATORY DEVELOPMENT	ADVANCED DEVELOPMENT
ACTIVE IR (3-D)	1	1	0	2
PASSIVE IR	0	4	0	1
MM	3	0	0	1
FIXED TARGET RADAR	3	1	1	5
MOVING TARGET RADAR	3	0	0	0

CONCLUSION

Table III summarizes the efforts ongoing or planned by the Reconnaissance and Weapon Delivery Division of the Avionics Laboratory. The exploratory development efforts are attempting to prove the utility of the fundamental technology. The advanced development activity evaluates the technology in high performance military environment. Table III indicates trends in the state of maturity of the various technologies. Efforts in processing passive infrared imagery are the most mature.

The processing technology is key to allowing pilots in single seat aircraft to deal with the target acquisition task loading. In a target rich environment, pilots need to know which targets are threats, especially in a defense suppression mission. The improved automatic screening performance is obtained through the integration of improved target acquisition sensors and associated processing. The integrated screener technology offers the following improvements:

- Reduced aircrew workload

- Automatic nomination of correct targets

- More survivable flight profiles

- Correct weapons against each target

- Targeting decisions at longer ranges

Rapid mobility of modern military forces dictate two fundamental requirements for reconnaissance operations: timeliness and completeness. Timely targeting information is needed for command decisions and reconnaissance operations are incomplete unless information is accurate and available during day/night and all weather conditions. Timeliness and completeness are opposing attributes in that rapid access implies a limited quantity of highly selective data, whereas comprehensive coverage under diverse conditions, with a high probability of success, implies a large amount of data. The approach being taken by the Avionics Laboratory to reduce this dichotomy is to employ technological advancements in sensor performance, automatic data processing, automatic cueing and classification sensors, and effective correlation and data handling techniques. In addition, a sensor or processing concept is not considered viable unless it is affordable and compatible with operational systems and concepts.

REFERENCES

Gardiner, F. J., Nicholson, W. E., April 1971, Data Management for Improved Tactical Reconnaissance and Multi-Sensor Applications. Proceedings for the Symposium on Tactical Reconnaissance and Surveillance, Institute for Defense Analysis, Arlington, VA.

Van Keuk, G., 1979, "Processing of Airborne Reconnaissance Data for In-Flight Display and Near Real-Time Transmission", AGARD-AR-135.

APPENDIX A

LIST OF ATTENDEES

ADAMS, F.N. Mr — Marconi Avionics Ltd., Elstree Way, Borehamwood, Herts., UK

ALDERS, G.J. Ir. — National Aerospace Laboratory, Anthony Fokkerweg 2, 1059 CM Amsterdam, NE

BAARS, E.P. Mr — FGAN/FHF, Koenigstrasse 2, 5307 Wachtberg-Werthoven, GE

BAFFERT-FORGES, G. Mr — Societe Matra, 37 Ave Louis Breguet, 78140 Velizy, FR

BALL, W.F. Mr — Head, Avionic Facilities Division, Commander, Naval Weapons Center, Code 3111, China Lake, CA 93555, US

BALSTON, D. Dr — Plessey Co. Ltd., Electronic Systems Research, Southleish Park House, Havant, Hampshire, UK

BECKMANN, L.H.J.F. Dr — Adjunkt Direkteur S.P.O., NV Optische Industrie de Oude Delft, Postbus 72, 2600 MD Delft, PAYS-BAS, GE

BERRY, T.R. Mr — G Blk, RSRE St. Andrews Road, Malvern, UK

BETEMPS, M. Mr — Aeritalia-B.A.I.P.A., Groupo Equipaggiamenti, Esercizio di Caselle, I-10072 Caselle-Turin, IT

BINNS, C.W. Mr — British Aerospace, Dynamics Group, Bristol Works, Filton, Bristol, UK

BIRD, P.A. Miss — British Aerospace, Dynamics Group IRE Div., Manor Road, Hatfield, Hertfordshire, AL10 9LL, UK

BLAKE, R.E. Mr — Computing Science Branch, Royal Military College of Science, Shrivenham, Swindon, Wilts, UK

BODIANSKY, M. Mr — ONERA, 92320 Chatillon, FR

BOHNER, M. Dr Ing — FIM, Breslauer Strasse 48, 7500 Karlsruhe 1, GE

BORN, G.K.A. Dr — MBB, AE101, P.O.Box 801149, D-8000 Munchen 80, GE

BOSMAN, D. Prof. Ir — Blds. EF, Twente University, P.O.Box 217, 7500 Enschede, NE

BOUDAREL, R. Mr — CIMSA, 10–12 Ave de L'Europe, 78140 Velizy, FR

BOUIGE, G. Mr — Thomson-CSF, 52 rue Guynemer, 92130 Issy-les-Moulineaux, FR

BRAULT, Y. Mr — Thomson-CSF Div. Equipementes, Avioniques & Spatiaux, 178 Bld Gabriel Peri, 92240 Malakoff, FR

BROCK-NANNESTAD, L.V. Dr — Danish Defence Research Establishment, Staunings Plads 2, 2100 Copenhagen $\emptyset$, DE

BROGI, G. Dr — c/o Sistel, via Tiburtina, 1210, 00131 Rome, IT

BROOKS, S.R. Dr — Marconi Research Labs, West Hanningfield Road, Great Baddow, Chelmsford, Essex CM2 8HN, UK

BROWN, G. Dr — Marconi Avionics, EDASD, D Building, Paycocke Rd, Basildon, Essex, UK

BUGGE-ASPERHEIM, P. Mr — NDRE, NO

BURKE, N.G. Dr — Smith Associates, 45–47 High Street, Cobham, Surrey, UK

CALLEWAERT, J. Mr — Air Staff Brussels VDT/A, Everestraat, 1140 Brussels, BE

CAREW-JONES, R.D. Mr — Computing Devices Co. Ltd., P.O.Box 10, Castleham Rd., St. Leonards-on-Sea, East Sussex, UK

CHABAH, M. Mr — Thomson-CSF, 178 Bld Gabriel Peri, 92240 Malakoff, FR

CONNAN, C. Mr — Avions Marcel Dassault, 78 Quai Carnot, 92214 St. Cloud, FR

CONRADY, T. Mr — Dornier Gmbh, Abt. AS 20, Postfach 1420, D7990-Friedrichshafen, GE

COOPER, C.W. Mr — RSRE, R102, St. Andrews Rd, Malvern WR14 3PS, UK

CORNET, J. Mr — Thomson-CSF, 52 rue Guynemer, 92130 Issy-les-Moulineaux, FR

CRAMIER, A. Mr — Thomson-CSF, 178 Bld Gabriel Peri, 92240 Malakoff, FR

DETTWILER, W. Mr — MacDonald, Dettwiler & Associates Ltd., 10280 Shellbridge Way, Richmond, B.C. V6X 2Z9, CA

DEY, D. Dr	Vereinigte Flugtechnische Werke Gmbh, Hunefeldstrasse 1–5, 2800 Bremen 1, GE
DIAMOND, F.I. Mr	RADC/DC, Griffiss AFB, NY 13441, US
DIECKMANN, H. Mr	Messerschmitt Bolkow Blohm, UB Apparate/AE144, Postfach 801149, D-8000 Munchen 80, GE
DITTEL, R.H. Mr	DFVLR, I 554/D, D-8301 Wessling, GE
DORTOMB, B. Mr	Electronique Marcel Dassault, 55 Quai Carnot, 92210 St. Cloud, FR
DOVE, B.L. Mr	Head, Avionics Systems Branch, Electronics Directorate, M.S.477, NASA Langley Research Ctr., Hampton, VA 23665, US
DUKE, D.B. Mr	Infra Red Equipment Div, British Aerospace Dynamics Gp., Manor Rd.. Hatfield, Herts AL10 9LL, UK
EIBERT, M. Dr	Dornier Gmbh, Postfach 1420, Abt. BS-30, 7990 Friedrichshafen 1, GE
EKRE, H. Mr	A/S Kongsberg Vapenfabrikk, c/o NDRE, P.O.Box 25, N-2007 Kjeller, NO
ENTZMINGER, J. Mr	RADC/DCI, Griffiss AFB, NY 13441, US
ENZINGER, F.J. Mr	Messerschmitt Bolkow Blohm, UE Flugzeuge/FE73, Postfach 801160, D-8000 Munchen 80, GE
ERNST, V. Dr	Bodenseewerk Geraetetechnik, P.O.B.1120, 7770 Uberlingen, GE
ESSEN, H. Dr	FGAN/FHP, Koenigstr.2, 5307 Wachtsberg-Werthoven, GE
EVANS, D.W. Mr	Avionics Systems Dept, British Aerospace Ltd., Richmond Road, Kingston-upon-Thames, Surrey, UK
FAULKNER, B.M. Mrs	MBB Gmbh, UB Drehflugler/DE 235, Postfach 801140, 8000 Munchen 80, GE
FOYE, J.M. Mr	Short Bros. Ltd., Missile Systems Div., Montgomery Road, Castlereagh, Belfast BT6 9HN, UK
FRIESE, D. Mr	Messerschmitt Bolkow Blohm, UB Apparate/AE411, Postfach 801149, D-8000 Munchen 80, GE
GALLO, G. Mr	Ufficiale E.I., Camen S. Piedro, A Grado, Pisa, IT
GEIGER, Karl Mr	Messerschmitt Bolkow Blohm, UB Apparate/AE 212, Postfach 801149, D-8000 Munchen 80, GE
GERHARDSEN, T.B. Mr	A/S Kongsberg Vapenfabrikk, Postbous 25, N-3601 Kongsberg, NO
GERHARDT, L.A. Prof.	Rensseleer Polytechnic Institute, Dept. of Electrical Eng., Troy, N.Y.12181, US
GHICOPOULOS, B. Mr	Communications Dept. Hellenic Air Fo., Technology Research Ctr., (KETA) Delta Falirou, P.Faliron, Athens, GR
GHIONNA, C. Mr	Contraves Italiana, via Affile 102, 00131-Roma, IT
GINET, C. Mr	T.R.T., DTM/OPT 5 Av. Reaumur, 92350 Plessis-Robinson, FR
GIRARD, M. Mr	CTRE/OP, ETCA, 94114 Arcueil CEDEX, FR
GRICE, J. Mr	Future Avionics Systems, EASAMS Ltd., Frimley Road, Frimley, Camberley, Surrey, UK
GROSCH, G. Mr	Messerschmitt Bolkow Blohm, UB Flugzeuge/FE 301, Postfach 801160, D-8000 Munchen 80, GE
GUDMANDSEN, P.E. Prof.	Laboratory-Electromagnetic Theory, Technical University, Bldg 348, Lundtoftevej 100, 2800 Lyngby, DE
GUILDFORD, L.H. Mr	Philips Research Labs., Cross Oak Lane, Redhill, Surrey RH1 5HA, UK
HALLDORSSON, T. Mr	Messerschmitt Bolkow Blohm, UB Apparate/AE 202, Postfach 801149, D-8000 Munchen 80, IC
HALLIWELL, D. Mr	Decca SSM Division, Cox Lane, Chessington, Surrey KT9 1SB, UK
HANSON, V.S.G. Mr	AEE, RISO, AMC, RDAF, Postbox 130, DK3500 Vaerloese, DE
HARP, J.G. Mr	RSRE, St. Andrews Road, Malvern, Worcs, UK
HASLAM, G.E. Mr	Dept. of Communications, Communications Research Centre, Box 11490 Station H, Ottawa, K2K 852, CA
HEDEN, S. Mr	Danish Army Materiel Command, Radar Branch, 55, Arsenalves, DK-9800 Hjorring, DE
HENDERSON, C.M. Mr	RSRE, St. Andrews Road, Malvern, Worcs, UK
HERION, H.J. Mr	Messerschmitt Bolkow Blohm, UB Apparate/AE 145, Postfach 801149, D-8000 Munchen 80, GE
HESSE, K. Dr	Herrn Johannes Hohnen, AEG-Telefunken, Industriestr. 31, 2000 Wedel/Holstein, GE
HIRZINGER, G. Dr	DFVLR-FF-DF, D-8031 Wessling, GE
HOFFMANN, H.E. Mr	DFVLR-NE-PA, 8031 Wessling-Oberpfaffenhofen, GE

HORNFELD, W. Mr	VFW, Huenefeldstr. 1 5, 2800 Bremen, GE
HUGHES, G.R. Mr	ROME Air Dev. Ctr., Attn. IRR, G.Hughes, Griffiss AFB, NY 13411, US
HUISMAN, W.C. Mr	National Aerospace Laboratory, Anthony Fokkerweg 2, Amsterdam, NE
ISCHEN, H. Dipl. Ing	F.I.M., Breslauer Strasse 46, 7500 Karlsruhe 1, GE
JACK, J.W. Dr	Electro-Optics Gp., Ferranti Ltd., Robertson Ave, Edinburgh, Scotland, UK
JACKSON, R.M.E. Mr	EMI Electronics Ltd., Victoria Road, Feltham, Middlesex, UK
JACOBSEN, M. Mr	AEG-Telefunken N14 V3, Postfach 1730, D-7900 Ulm, GE
JAMES, E. Mr	Systems Control Dept., Site A, British Aerospace, Dynamics Group, Stevenage, Herts, UK
JANTZEN, R. Mr	BMVG, Forschinstitut fur Optik, Schloss Kressbach, D-7400 Tubingen 1, GE
JENSEN, G.H. Mr	Danish Defence Research Est., Staunings Plads 2, 2100 Copenhagen Ø, DE
JONES, T.L. Mr	Night Vision & Electro-Optics Lab., Advanced Concepts Division, Target Acq. Res. Team, Ft. Belvoir, VA 22060, US
KAESTNER, R. Mr	VFW Bremen, Huenefeldstr. 1–5, 2800 Bremen, GE
KELLER, E. Mr	Messerschmitt Bolkow Blohm, UB Apparate/AE 145, Postfach 801149, D-8000 Munchen 80, GE
KIERKEGAARD, F. Mr	AMC, RDAF, Postbox 130, DK 3500 Vaerloese, DE
KISTER, H. Mr	VDO Luftfahrtgerate Werk, An der Sandelmuhle 13, GE
KJEMHUS, K. Mr	A/S Kongsberg Vapenfabrikk, Postbox 25, N-3601 Kongsberg, NO
KNOCH, M.H. L/Col.	HQ Royal Netherlands Air Force, P.O.Box 90501 LM, The Hague, NE
KOHNLE, A. Dr	Forschungsinstitut fur Optik, Schloss Kressbach, D-7400 Tubingen 1, GE
KOSTKA, M.E. Mr	EMI Electronics Ltd., 4th Floor CF, Blyth Road, Hayes. Middx, UK
KRAUTWALD, W. Dr	Messerschmitt Bolkow Blohm, UB Flugzeuge/FE 324, Postfach 801660, D-8000 Munchen 80, GE
KUNZMANN, R.G. Mr	ELTRO Gmbh, Abteilung V1, Postfach 102120, 6900 Heidelberg 1, GE
LAIER, W. Mr	Messerschmidt-Bolkow-Blohm Gmbh, Zentralbar/ZD 226, Postfach 801149, 8000 Munchen 80, GE
LAMATSCH, H. Dr	Messerschmitt Bolkow Blohm, UB Flugzeuge/FE, Postfach 860629, D-8000 Munchen 80, GE
LAFP, H.S. Mr	AFWAL/AARI, Wright-Patterson AFB, Ohio, 45433, US
LE CHEVALIER, F. Mr	ONERA, 92320 Chatillon, FR
LECLUYSE, J. Lt.	Ave. A. Lacomble, 17, B.1140 Brussels, BE
LEVERS, R.G. Dr	DR Division, AUWE(N), Portland, Dorset, UK
LEWIS, T.A. Mr	Marconi Avionics, Electro-Optic Advanced Systems, D Bldg., Paycocke Road, Basildon, Essex, UK
LLOYD, D.G. Mr	EMI Electronics, Victoria Road, Feltham, Middlesex, UK
LORENZONI, A. L/Col.	Aeronautica Militaire, Centro Consultivo Studi de Recherche, via DEI Pontefici, 3, 00186 Rome, IT
LOUCHET, J. Mr	ETCA, 16 bis Ave Prieur de la Cote D'Or, 94114 Arcueil-CEDEX, FR
LUDBROOK, F.T. Mr	Plessey Co. Ltd., Electronic Systems Research, Southleigh Park House, Havant, Hampshire PO9 ZPE, UK
LUEG, H. Prof.	Institut fur Technische Elektronik, RWTH, Templegraben 55, D-5100 Aachen, GE
MACMEEKIN, N. Ms.	Naval Air Development Center, Code 3011, Warminster, PA 18974, US
MASCARENHAS, J.M.B.G. Capt.	Rua da Escola do Exercito, 13, 1100 Lisboa, PO
MAVKO, G.E. Dr	General Dynamics Corp., Convair Division MZ 42-6210, P.O.Box 80647, San Diego, CA 92038, US
MAYES, D.J. Mr	Smith's Industries Aerospace and Defence Systems Co., Cheltenham Div., Bishop's Cleeve, Cheltenham, Glos GL52 4SF, UK
MENKE, J.F. Mr	Elektro-Optik Gmbh & Co. KG, Fordestr. 35, 2392 Glucksburg, GE
MILLER, D. Mr	Avionics Development Dept., 7 Hangar, Warton Aerodrome, Warton, Lancs, UK
MONTERMANN, E.J. Mr	Messerschmitt Bolkow Blohm, UB Apparate/AE 411, Postfach 801149, D-8000 Munchen 80, GE
MOREAU, C. Ica	STTE/PNI, 129 rue de la Convention, 75731 Paris CEDEX 15, FR
MUIR, N. Mr	B. AeD. Elect. Opt. Dept. P.B.244, Site A, Six Hills Way, Stevenage, Herts, SG1 2DA, UK

MULLER, M. Mr	ESG, Vogelweidesplatz 9, Postfach 800569, D-8000 Munchen 80, GE
McKENDRICK, W.H. Dr	Ferranti Ltd., Electro Optics Gp., St. Andrews Works, Robertson A, Edinburgh, UK
McKINLAY, W.H. Mr	Ferranti Ltd., Silverknowes, Edinburgh EH4 4AD, Scotland, UK
NINZ, N.R. Dr	VFW Gmbh, Hunefeldstrasse 1−5, 2800 Bremen 1, GE
NORDGARD, K. Mr	A/S Kongsberg Vapenbaerikk, Postbox 25, N-3601 Kongsberg, NO
OLBERTZ, A.H. Mr	Hollandse Signaalaparten BV, Zuid. Havenweg 40, Hengelo, NE
OSTEVOLD, E. Mr	NDRE, Postbox 25, 2007 Kjeller, NO
PADDEN, W.J. Mr	Dept. Trade & Commerce, Transportation Industries Branch (53), 6th Floor E., 235 Queen Street, Ottawa, CA
PAUL, D. Mr	IITB der Fraunhofer-Geselschaft, Sebastian-Kneippstr. 12, D-7500 Karlsruhe 1, GE
PEEBLES, K.A. Mr	Defence Research Establishment, National Defence Headquarters, Ottawa, Ontario KIA OZ4, CA
PINKNEY, H.F.L. Dr	National Aeronautical Establishment, National Research Council Canada, Ottawa, Ontario K1A 0R6, CA
POWELL, J.M. Mr	British Aerospace Aircraft Gp., Warton Division, Preston, Lancaster, UK
PRINCE, J.R. Mr	MOD (PE), NAV/C, Room 403, St. Giles Ct., 1−13 St. Giles High St., London WC2H8LD, UK
REBUFFET, M. Mr	CTRE/OP ETCA, 94114 Arcueil CEDEX, FR
RHODES, S.L. Mr	RACAL-SSA Ltd., Basingstoke Rd., Spencers Wd, Reading RG7 1PD, UK
RIDLER, T.W. Dr	Attack Weapons Dept. Y20, Royal Aircraft Est., Farnborough, Hants, UK
RIEUS, F. Mme	Electroniqe Marcel Dassault, 55 Quai Carnot, 92214 St. Cloud, FR
ROGGE, J. Mr	Hoge Zijde 32, 4584 AH Bavel, NE
ROY, S.D. Mr	Head, Advanced Systems, Westland Helicopters Ltd., Keymarket House, Yeovil, Somerset, UK
RUECKERT, P. Dr	in IABG Gmbh, Einsteinstr, D-8012 Ottobrunn, GE
RYLES, J.C. Dr	Chief Scientist, AFWAL/AS, Wright-Patterson AFB, Ohio 45433, US
SALKELD, G. Mr	c/o British Aerospace Aircraft Div., Warton Aerodrome, Warton, Preston, Lancs, UK
SCHIEBSCHICK, L. Mr	BMVG RU I 5, Sibe, 5300 Bonn 1, GE
SCHMIDT, E. Mr	VFW Gmbh, Huenefeldstr. 1−5, 2800 Bremen, GE
SCHOLLER, P.H. Mr	Messerschmitt Bolkow Blohm, UB Apparate/AE 214, Postfach 801149, D-8000 Munchen 80, GE
SCHOENE, G. Mr	Herrn Johannes Hohnen-AEG-Telefunken, Industriestr. 31, D-2000 Wedel/Holst, GE
SCHRANNER, R. Dr	Messerschmitt Bolkow Blohm, UB Flugzeuge/FE 414, Postfach 801160, D-8000 Munchen 80, GE
SCHRODER-BRZOSNIOWSKY, J. Dr	Dornier Gmbh, Abt. BS-10, Postfach 1420, 7990 Friedrichshafen, GE
SCHWANDTNER, K. Dr	Battelle-Institute, Am Romerhof 35, 6000 Frankfurt/MAIN 90, GE
SEIFERTH, R. Mr	Siemens AG, Abt. WV Fu Ort 3, 70 Hofmannstr. 51, 8000 Munchen 70, GE
SIRVEN, J. Mr	Electronique Marcel Dassault, 55 Quai Carnot, 92214 St. Cloud, FR
SOLAND, D.E. Mr	Honeywell, Inc, 2600 Ridgway Parkway, Minneapolis, Minn 55413, US
STEINHARDT, R. Mr	AEG Telefunken, Industrietstr. 31, D-2000 Wedel/Holstein, GE
STENGEL, G.G. Dipl. Ing	ELTRO Gmbh, Postfach 102120, D-6900 Heidelberg 1, GE
STENNER, K.H. Mr	Messerschmitt Bolkow Blohm, UB Drehflingler/DE 112, Postfach 801120, D-8000 Munchen 80, GE
STIES, M. Dr-Ing	F.I.M., Breslauer Strasse 48, 7500 Karlsruhe 1, GE
STRINGER, F.S. Mr	ADXR/FS RAE/UK, Flight Systems Dept., Royal Aircraft Est., R177 Bldg. Farnborough, Hants, UK
STROKE, G.W. Prof. Dr	Messerschmitt Bolkow Blohm, Zentralber/ZET, Postfach 801109, 8000 Munchen 80, GE
STROMMAN, E. Mr	NDRE, P.O.Box 25, N-2007 Kjeller, NO
SUETA, T.J. Mr	U.S. Army Avionics R&D Activity, Attn. DAVAA-D, Ft.Monmouth, NJ 07703, US
TAILLET, J. Dr	Directeur Scientifique, ONERA, 92320 Chatillon, FR
TAYLOR, D.J. Mr	E.O.G., Ferranti Ltd., Robertson Ave, Edinburgh EH11 1PX, UK
TAYLOR, D.G. Mr	MEL, Manor Royal, Crawley, West Sussex, RH10 2PZ, UK

THERBURG, R.D. Dipl. Phys.	c/o Prof. Muhlenfeld, Lehrstuhl fur Regeltechnik und Elektronik der tu Clausthal, Leibnizstr. 28, 3392 Clausthal-Zellerfeld, GE
TIMMERS, H.A. Ir.	National Aerospace Research Laboratory, Anthony Fokkerweg 2, 1006 BM Amsterdam, NE
VALETON, J.M. Mr	Institute for Perception TNO, Kampweg 5, 3769 DE Soesterberg, NE
VAN KASTEEL, J. Mr	National Aerospace Research Laboratory, Anthony Fokkerweg 2, Amsterdam, NE
VLEESHDRAAGER, E.N. Ir.	N.V. Optische Industrie, De Oude Delft, Postbus 72, 2600 MD Delft, NE
VOGEL, M. Dr Ing	DFVLR e.v. 8031 Oberpfaffenhofen, Post Wessling/OBB, GE
VOGEL, H.H.E. Mr	Danish Defence Research Est., Staunings Plads 2, 2100 Copenhagen Ø, DE
VOLES, R. Mr	EMI Electronic Ltd., 135 Blythe Road, Hayes, Middlesex, UB3 1BP, UK
WEIMANN, A. Dr	ESG Elektronik System Gmbh, Postfach 800569, Vogelweideplatz 9, 8000 Munchen 80, GE
WEISS, M.T. Dr	The Aerospace Corporation, P.O.Box 92957, Los Angeles, CA 90009, US
WESLEY, A.C. Mr	Ferranti Ltd., Ferry Road, Edinburgh EH5 2XS, Scotland, UK
WICHMANN, G. Mr	ELTRO Gmbh, D-6900 Heidelberg, Kurpfalzring 106, GE
WIEMER, W. Dr	Messerschmitt Bolkow Blohm, UB Apparate/AE 145, Postfach 801149, D-8000 Munchen 80, GE
WIEMKEN, U. Dr	I.N.T./FhG, Appelsgarten 2, D-535 Euskirchen, GE
WILD, J. Mr	Messerschmitt Bolkow Blohm, UB Flugzeuge/FE 117, Postfach 801160, D-8000 Munchen 80, GE
WILLIAMS, J.S. Mr	British Aerospace, Aircraft Gp., Hqs., Richmond Road, Kingston-upon-Thames, Surrey, KT2 5QS
WOLF, W.E. Mr	RADC/ECTM, Griffiss AFB, N.Y. 13441, US
YNDESTAD, H. Mr	NDRE, P.O.Box 25, N-2007, Kjeller, NO
ZEIFANG, G. Dr	Elektro-Optik Gmbh, Fordestr. 35, D2392 Glucksburg, GE
ZURAN, J. Mr	c/o Technology Branch, Dept. Industry Trade & Commerce, 235 Queen St., 8-E Flr., Ottawa, Ontario K1A 0H5, CA

REPORT DOCUMENTATION PAGE

1. Recipient's Reference	2. Originator's Reference	3. Further Reference	4. Security Classification of Document
	AGARD-CP-290	ISBN 92-835-0280-9	UNCLASSIFIED

5. Originator	Advisory Group for Aerospace Research and Development North Atlantic Treaty Organization 7 rue Ancelle, 92200 Neuilly sur Seine, France

6. Title	IMAGE AND SENSOR DATA PROCESSING FOR TARGET ACQUISITION AND RECOGNITION

7. Presented at	the 40th Technical Meeting, of the Avionics Panel of AGARD, held in Aalborg, Denmark on 8–12 September, 1980.

8. Author(s)/Editor(s)	9. Date
Various	November 1980

10. Author's/Editor's Address	11. Pages
Various	274

12. Distribution Statement	This document is distributed in accordance with AGARD policies and regulations, which are outlined on the Outside Back Covers of all AGARD publications.

13. Keywords/Descriptors

Target acquisition	Man machine systems
Target recognition	Tracking (position)
Data processing	Visibility
Images	

14. Abstract

These Proceedings contain the Unclassified papers presented at the 40th Technical Meeting of the Avionics Panel of AGARD, which was held in Aalborg, Denmark, 8–12 September, 1980. They consist of 21 papers on the subject of "Image and Sensor Data Processing for Target Acquisition and Recognition", which are divided as follows: target backgrounds and visibility, 3; man-machine interaction, 2; image processing, 6; target tracking, 4; target classification and identification, 6. The companion volume which is classified Nato-Secret contains the remaining 17 papers which were classified and the discussions and summary. If possible both texts should be reviewed simultaneously.

AGARD Conference Proceedings No.290
Advisory Group for Aerospace Research and Development, NATO
IMAGE AND SENSOR DATA PROCESSING FOR TARGET ACQUISITION AND RECOGNITION
Published November 1980
274 pages

These Proceedings contain the Unclassified papers presented at the 40th Technical Meeting of the Avionics Panel of AGARD, which was held in Aalborg, Denmark, 8–12 September, 1980. They consist of 21 papers on the subject of "Image and Sensor Data Processing for Target Acquisition and Recognition", which are divided as follows: target backgrounds and visibility, 3; man-machine interaction, 2; image processing, 6; target

P.T.O.

AGARD-CP-290

Target acquisition
Target recognition
Data processing
Images
Man machine systems
Tracking (position)
Visibility

---

tracking, 4; target classification and identification, 6. The companion volume which is classified Nato-Secret contains the remaining 17 papers which were classified and the discussions and summary. If possible both texts should be reviewed simultaneously.

ISBN 92-835-0280-9

tracking, 4; target classification and identification, 6. The companion volume which is classified Nato-Secret contains the remaining 17 papers which were classified and the discussions and summary. If possible both texts should be reviewed simultaneously.

ISBN 92-835-0280-9

tracking, 4; target classification and identification, 6. The companion volume which is classified Nato-Secret contains the remaining 17 papers which were classified and the discussions and summary. If possible both texts should be reviewed simultaneously.

ISBN 92-835-0280-9

tracking, 4; target classification and identification, 6. The companion volume which is classified Nato-Secret contains the remaining 17 papers which were classified and the discussions and summary. If possible both texts should be reviewed simultaneously.

ISBN 92-83ʹ-0280-9